

# Quantum Algorithms Using the Curvelet Transform

Yi-Kai Liu

Institute for Quantum Information  
California Institute of Technology  
Pasadena, CA, USA  
yikailiu@caltech.edu

Mar. 25, 2009

## Abstract

The curvelet transform is a directional wavelet transform over  $\mathbb{R}^n$ , which is used to analyze functions that have singularities along smooth surfaces (Candès and Donoho, 2002). I demonstrate how this can lead to new quantum algorithms. I give an efficient implementation of a quantum curvelet transform, together with two applications: a single-shot measurement procedure for approximately finding the center of a ball in  $\mathbb{R}^n$ , given a quantum-sample over the ball; and, a quantum algorithm for finding the center of a radial function over  $\mathbb{R}^n$ , given oracle access to the function. I conjecture that these algorithms succeed with constant probability, using one quantum-sample and  $O(1)$  oracle queries, respectively, independent of the dimension  $n$  — this can be interpreted as a quantum speed-up. To support this conjecture, I prove rigorous bounds on the distribution of probability mass for the continuous curvelet transform. This shows that the above algorithms work in an idealized “continuous” model.

## 1 Introduction

One of the most remarkable demonstrations of the power of a quantum computer is Shor’s algorithm for factoring and discrete logarithms [26]. This has motivated many researchers to try to generalize its key components—the quantum Fourier transform over  $\mathbb{Z}_N$ , and the algorithm for period-finding—to solve other problems [13]. In particular, this motivated the study of the quantum Fourier transform and the hidden subgroup problem (HSP) on non-Abelian groups, as a route to solving certain lattice problems and the graph isomorphism problem [23, 3, 25].

In this paper we study a different generalization of the Fourier transform, namely the curvelet transform over  $\mathbb{R}^n$  [10]. This is a kind of “directional” wavelet transform, which can resolve features over the spatial and frequency domains simultaneously. A curvelet basis function resembles a wavepacket, with high-frequency oscillations in one direction (like a plane wave  $e^{i\vec{k}\cdot\vec{x}}$ , as in the Fourier transform on  $\mathbb{R}^n$ ), but which is also supported on a small region of space (unlike the plane wave). We show that this leads to fast quantum algorithms for some new classes of problems, outside the framework of the HSP. To the best of our knowledge, this is the first attempt to design quantum algorithms based on the curvelet transform.

Intuitively, the curvelet transform is helpful in analyzing functions on  $\mathbb{R}^n$  that are discontinuous along  $(n-1)$ -dimensional surfaces. If a function  $f$  is discontinuous along a surface  $S$ , then its curvelet transform  $\Gamma_f$  will be “large” at those locations  $(\vec{b}, \vec{\theta})$ , where  $\vec{b}$  is a point on  $S$  and  $\vec{\theta}$  is the vector normal to  $S$  at  $\vec{b}$ . The set of all such pairs  $(\vec{b}, \vec{\theta})$  is called the “wavefront set” of  $f$ .

The basic model for a quantum algorithm using the curvelet transform is as follows: first prepare a quantum state  $\sum_{\vec{x} \in \mathbb{R}^n} f(\vec{x})|\vec{x}\rangle$ , which is a weighted superposition of points in  $\mathbb{R}^n$ ; then apply the quantum curvelet transform, to get the state  $\sum_{\vec{b}, \vec{\theta}} \Gamma_f(\vec{b}, \vec{\theta})|\vec{b}, \vec{\theta}\rangle$ ; and finally measure  $\vec{b}$  and  $\vec{\theta}$ . For the time being, we ignore implementation issues, such as how to discretize  $\mathbb{R}^n$  and how to compute the quantum curvelet transform efficiently. The more basic question is whether we can find functions  $f$  such that we can prepare the initial state efficiently, and such that a measurement in the “curvelet basis” yields useful information.

One example consists of letting  $f$  be the indicator function of a ball in  $\mathbb{R}^n$ . We can efficiently prepare a uniform superposition over points in a ball, using the techniques of [2, 17]. (This is called “quantum-sampling,” and it can be done more generally, e.g., for convex bodies.) Then, measuring in the curvelet basis extracts information about the center of the ball.

Another example consists of choosing  $f$  to be the indicator function of a spherical shell in  $\mathbb{R}^n$ . This is motivated by the problem of finding the center of a radial function on  $\mathbb{R}^n$ . Let  $G$  be a radial function, centered around some unknown point  $\vec{c}$ . Prepare a uniform superposition over a large region of space, then compute  $G(\vec{x})$  and measure it; this produces a uniform superposition over one of the level sets of  $G$ , which is a spherical shell centered at  $\vec{c}$ . Then, measuring in the curvelet basis extracts information about the location of  $\vec{c}$ .

The goal of this paper is to make this intuition precise. We can interpret  $\Gamma_f$  as a wavefunction, and we want to show that its probability mass  $|\Gamma_f|^2$  is concentrated near the wavefront set. We can prove this for the continuous curvelet transform, for the two cases of interest, where  $f$  is the indicator function of a ball or a spherical shell in  $\mathbb{R}^n$ . In these cases,  $(\vec{b}, \vec{\theta})$  is near the wavefront set with high probability. This implies that the line  $\{\vec{b} + \lambda \vec{\theta} \mid \lambda \in \mathbb{R}\}$  passes near the center of the ball or spherical shell.

Next, we give an efficient implementation of a quantum curvelet transform. (This is a discrete version of the transform described above, acting on superposition states.) Then we propose polynomial-time quantum algorithms for two problems: (1) given a single quantum-sample over a ball in  $\mathbb{R}^n$ , find the center of the ball, with accuracy  $\pm\Delta$  where  $\Delta$  is a constant fraction of the radius of the ball; (2) given oracle access to a radial function  $f$  that is centered around some unknown point  $\vec{c} \in \mathbb{R}^n$ , find the point  $\vec{c}$  exactly (i.e., with accuracy  $\pm\Delta$  in time  $\text{poly}(\log \frac{1}{\Delta})$ , assuming that the function  $f$  fluctuates on sufficiently small scales).

For the first problem, we conjecture that our quantum procedure succeeds with constant probability, while the best classical procedure succeeds with probability that is exponentially small in  $n$ . Classically, this problem is hard because the volume of a ball in  $\mathbb{R}^n$  is concentrated near its surface. But this same fact is helpful to the quantum curvelet transform, which works by finding a line normal to the surface of the ball.

For the second problem, we conjecture that our quantum algorithm uses only a constant number of queries, whereas any classical algorithm requires  $\Omega(n)$  queries. Intuitively, this is because the curvelet transform uses constructive interference to find a direction in  $\mathbb{R}^n$  from just one query.

We then prove that these algorithms work in an idealized “continuous” model — this follows from our rigorous results on the continuous curvelet transform. However, we do not have a rigorous analysis of the effects caused by discretization, though we can argue that these should be small.

These examples demonstrate that one can use the curvelet transform to obtain a quantum speed-up. These examples are artificially simple, in order to allow a rigorous analysis. But the underlying idea—using the curvelet transform to find normal vectors to a surface—should work on more complicated geometric objects.

**1.1 Technical Contributions:** First, in Section 2, we define the continuous curvelet transform over

$\mathbb{R}^n$ . This generalizes the definition over  $\mathbb{R}^2$  given in [8]. Given a function  $f(\vec{x})$ , the continuous curvelet transform returns a function  $\Gamma_f(a, \vec{b}, \vec{\theta})$ . Here,  $\vec{x} \in \mathbb{R}^n$  represents a “location,” while  $0 < a < 1$  is a “scale” (smaller values denote finer scales, larger values denote coarser scales),  $\vec{b} \in \mathbb{R}^n$  is a “location,” and  $\vec{\theta} \in S^{n-1}$  (the unit sphere in  $\mathbb{R}^n$ ) is a “direction.”

Next, we study the distribution of probability mass  $|\Gamma_f|^2$  over different values of  $(a, \vec{b}, \vec{\theta})$ . This is technically quite difficult.  $\Gamma_f(a, \vec{b}, \vec{\theta})$  is defined by an oscillatory integral, and while there are various methods for bounding the asymptotic decay rates of such quantities [28, 7], we need non-asymptotic bounds on the total probability mass in a given region. In Section 3 we develop some tools for proving such bounds, in the case where  $f$  is a radial function. Then, in Sections 4 and 5, we specialize to the case where  $f$  is the indicator function of a ball or a spherical shell. Here, the analysis relies on powerful classical results that bound the oscillation and decay of Bessel functions [1, 30].

In Section 4 we let  $f$  be the indicator function of a ball in  $\mathbb{R}^n$ , with radius  $\beta$ , centered at the origin. We expect that, after applying the curvelet transform,  $\vec{b}$  and  $\vec{\theta}$  will be concentrated near the wavefront set of  $f$ : that is,  $\vec{b}$  will be concentrated near the line  $\{\lambda\vec{\theta} \mid \lambda \in \mathbb{R}\}$ , at about distance  $\beta$  from the origin. Furthermore, we expect that  $\vec{b}$  will become more tightly concentrated, the smaller the value of  $a$ . We show that this is essentially what happens. In particular, with constant probability, the distance from  $\vec{b}$  to the line  $\{\lambda\vec{\theta} \mid \lambda \in \mathbb{R}\}$  will be at most a constant fraction of  $\beta$ ; remarkably, this holds independent of the dimension  $n$ .

In Section 5, we let  $f$  be supported on a thin spherical shell, having radius  $\beta$  and thickness  $\delta \ll \beta$ . Here, after applying the curvelet transform, we get a qualitatively similar behavior of  $a$ ,  $\vec{b}$  and  $\vec{\theta}$ . Quantitatively, however, we find that we can observe much smaller scales  $a$ , on the order of  $\delta/\beta$ ; and thus we can find the center of the shell with much greater precision. Essentially, by making the shell extremely thin, we can find its center with arbitrarily high precision.

Finally, we turn to the discrete curvelet transform, and quantum algorithms. In Section 6 we give an efficient implementation of a quantum curvelet transform. This uses ideas from the fast classical curvelet transform [6]. However, there is a new complication in the quantum case: we need to prepare certain states which are superpositions of different scales and directions  $(a, \vec{\theta})$ . We design families of curvelets that allow this step to be performed efficiently, and that have similar analytic properties to the curvelets used in Sections 3-5.

In Section 7, we formally define the two problems mentioned earlier: estimating the center of a ball, given a single quantum-sample state; and finding the center of a radial function, given oracle access. We present quantum algorithms for these problems, and use our results from Sections 4 and 5 to prove that the algorithms work in a continuous model. We also sketch a classical lower bound for finding the center of a radial function.

This paper omits most of the proofs, due to lack of space. The proofs can be found in the full version [21]. Also, note that this paper contains some additional results there were not present in the first version of [21]. This paper contains an improved algorithm for finding the center of a radial function, and a classical lower bound for that problem.

**1.2 Related Work:** Curvelets over  $\mathbb{R}^2$  and  $\mathbb{R}^3$  have been studied as a tool for image processing and simulating wave propagation [10, 9, 6, 20]. The curvelet transform is also related to older ideas from harmonic analysis; see, e.g., Smith [27]. It can be viewed as an algorithmic implementation of a technique known as second dyadic decomposition [28].

There are a few rigorous results on the behavior of the curvelet transform which are similar in spirit to our work [10, 7, 8, 5]. These results apply to much broader classes of functions, but they are only known to hold over  $\mathbb{R}^2$  (or  $\mathbb{R}^3$  in some cases). Although one would expect them to generalize in some fashion to  $\mathbb{R}^n$ , it is perhaps surprising that the scaling with  $n$  is as favorable as we find here.

In connection with quantum algorithms, there has been some work on the quantum wavelet transform [31, 15, 32, 16]. But the curvelet transform on  $\mathbb{R}^n$  is quite different from the “ordinary” wavelet transform on  $\mathbb{R}^n$ , which consists of a product of 1-D transforms. The ordinary wavelet transform on  $\mathbb{R}^n$  can detect the locations of discontinuities, but it cannot resolve directions.

The geometric problems studied in this paper are reminiscent of some recent work on finding hidden nonlinear structures, although the details are different. Shifted subset problems [12, 22] involve translational invariance, so the natural tool for solving them is the Fourier transform, rather than curvelets. Hidden polynomial problems [12, 14] resemble our problem of finding the center of a radial function. However, they are much more general (and thus much harder), and they are set over a finite field rather than  $\mathbb{R}^n$ .

These problems can also be studied from the perspective of quantum state discrimination [11], e.g., what is the optimal quantum measurement for estimating the center of a ball? However, we stress that our algorithms using the quantum curvelet transform are computationally efficient.

Finally, we recently became aware of a quantum algorithm for estimating the gradient of a function  $f$  on  $\mathbb{R}^n$ , using only  $O(1)$  queries [19]. This is quite similar to our algorithm for finding the center of a radial function — it is like applying the curvelet transform at a single location  $\vec{b}$ . Viewing this as a curvelet transform has the advantage of providing a more general framework, where one can do this procedure on an arbitrary input state. Also, note that our emphasis in this paper is on functions  $f$  that are not smooth—in this case, the gradient is not well-defined.

## 2 The Continuous Curvelet Transform

We begin by defining the continuous curvelet transform over  $\mathbb{R}^n$ . This generalizes the definition of [8] over  $\mathbb{R}^2$ . Given a function  $f(\vec{x})$ , the continuous curvelet transform returns a function  $\Gamma_f(a, \vec{b}, \vec{\theta})$ . Here,  $\vec{x} \in \mathbb{R}^n$  represents a “location,” while  $0 < a < 1$  is a “scale” (smaller values denote finer scales, larger values denote coarser scales),  $\vec{b} \in \mathbb{R}^n$  is a “location,” and  $\vec{\theta} \in S^{n-1}$  (the unit sphere in  $\mathbb{R}^n$ ) is a “direction.” All functions return values in  $\mathbb{C}$ .

Intuitively, the curvelet transform decomposes  $f$  into pieces corresponding to different scales  $a$  and directions  $\vec{\theta}$ ; we can view  $\Gamma_f(a, \vec{b}, \vec{\theta})$  as a family of functions, indexed by  $a$  and  $\vec{\theta}$ , each representing some “piece” of  $f(\vec{x})$  (note that the variables  $\vec{b}$  and  $\vec{x}$  both represent locations in space). To get one such piece of  $f$ , we will take its Fourier transform, and then multiply by a “window function”  $\chi_{a, \vec{\theta}}$  which is defined over the frequency domain.

More precisely, the curvelet transform is defined to be

$$\Gamma_f(a, \vec{b}, \vec{\theta}) := \int_{\mathbb{R}^n} \hat{f}(\vec{k}) \chi_{a, \vec{\theta}}(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{b}} d\vec{k}. \quad (1)$$

Here,  $\hat{f}$  is the Fourier transform of  $f$ , and  $\chi_{a, \vec{\theta}}$  is a function that is smooth, real, non-negative, and supported on a “sector” of frequency space  $S_{a, \vec{\theta}} \subset \mathbb{R}^n$ . We will describe these sectors below. Before doing so, we remark that the curvelet transform consists of (1) taking the Fourier transform of  $f$ , (2) separating  $\hat{f}$  into pieces corresponding to different scales and directions, and (3) taking the inverse Fourier transform. This description suggests how to compute the curvelet transform efficiently.

The “sector”  $S_{a, \vec{\theta}}$  is roughly given by the intersection of the cone centered around the vector  $\vec{\theta}$  with angular width  $\sqrt{a}$ , and the annulus with inner radius  $1/a$  and outer radius  $2/a$ . Thus, the “piece” of  $f$  at scale  $a$  and direction  $\vec{\theta}$  is somewhat like the restriction of  $\hat{f}$  to frequencies  $\vec{k} \approx (1/a)\vec{\theta}$  (which represent oscillations in direction  $\vec{\theta}$ , at higher frequencies when the scale  $a$  is small). Note,

however, that the sector has dimensions  $1/a \times (1/\sqrt{a})^{n-1}$ , so its shape is not constant — the sector becomes longer and narrower when the scale  $a$  is small.

This construction can also be understood from a second perspective. We can define a family of curvelet basis functions  $\gamma_{a,\vec{b},\vec{\theta}}$  as follows. The curvelet at location  $\vec{b}$  is obtained by translation from the curvelet at location  $\vec{0}$ , that is,  $\gamma_{a,\vec{b},\vec{\theta}}(\vec{x}) := \gamma_{a,\vec{0},\vec{\theta}}(\vec{x} - \vec{b})$ . The curvelet at location  $\vec{0}$  is defined in terms of its Fourier transform, which is simply the window function  $\chi_{a,\vec{\theta}}$ , that is,  $\hat{\gamma}_{a,\vec{0},\vec{\theta}}(\vec{k}) := \chi_{a,\vec{\theta}}(\vec{k})$ . It is easy to check that the curvelet transform defined earlier is equivalent to taking inner products with this family of curvelet basis functions:  $\Gamma_f(a, \vec{b}, \vec{\theta}) = \langle \gamma_{a,\vec{b},\vec{\theta}}, f \rangle$ .

Now we can see how our choice of the window function  $\chi_{a,\vec{\theta}}$  implies (and is motivated by) certain properties of the curvelet basis functions  $\gamma_{a,\vec{b},\vec{\theta}}$ . Since  $\chi_{a,\vec{\theta}}$  is smooth, the  $\gamma_{a,\vec{b},\vec{\theta}}$  are rapidly decaying. Also, each  $\gamma_{a,\vec{b},\vec{\theta}}$  has high-frequency oscillations in the  $\vec{\theta}$  direction, and is essentially supported on a plate-like region, centered at location  $\vec{b}$ , orthogonal to  $\vec{\theta}$ , with dimensions  $a \times (\sqrt{a})^{n-1}$ . Intuitively,  $\gamma_{a,\vec{b},\vec{\theta}}$  resembles a plane-wave in direction  $\vec{\theta}$ , localized around the point  $\vec{b}$ .

Finally, we define the window function  $\chi_{a,\vec{\theta}}$  as follows. Write  $\vec{k}$  using spherical coordinates  $(r, \phi_1, \dots, \phi_{n-1})$ , centered around the direction  $\vec{\theta}$ , so that  $\phi_1$  is the angle between  $\vec{k}$  and  $\vec{\theta}$ . Then define  $\chi_{a,\vec{\theta}}(\vec{k}) := W(\lambda ar) V(\phi_1/\sqrt{a}) \Lambda_a(\phi_1)$ . Here  $\lambda$  is a constant, which can be chosen freely; we will explain how to set it later.  $W$  is a radial window function, real, nonnegative, supported on the interval  $[1/e, 1]$ , and satisfying the admissibility condition  $\int_0^\infty W(r)^2 \frac{dr}{r} = 1$ .  $V$  is an angular window function, real, nonnegative, supported on the interval  $[0, \pi/2]$ , and satisfying the admissibility condition  $\int_{S^{n-1}} V(\phi_1)^2 d\sigma(\phi_1, \dots, \phi_{n-1}) = 1$ , where  $d\sigma$  denotes integration over the unit sphere  $S^{n-1}$  in  $\mathbb{R}^n$ .

$\Lambda_a$  is a normalization and adjustment factor:  $\Lambda_a(\phi_1) := a^{(n+1)/4} \left( \frac{\sin(\phi_1/\sqrt{a})\sqrt{a}}{\sin(\phi_1)} \right)^{(n-2)/2}$ . This is needed because the volume of the sector  $S_{a,\vec{\theta}}$  (on which  $\chi_{a,\vec{\theta}}$  is supported) changes with  $a$ . Note that, when  $\phi_1$  is small,  $\Lambda_a(\phi_1) \approx a^{(n+1)/4}$ . This is the main point where defining curvelets over  $\mathbb{R}^n$  is more complicated than over  $\mathbb{R}^2$ ; note that in dimension  $n = 2$ ,  $\Lambda_a(\phi_1) = a^{(n+1)/4} = a^{3/4}$  exactly. We remark that a simpler approach would be to use a constant normalization factor that only depends on  $a$  and not  $\phi_1$ ; however, our more complicated construction will be more convenient in the later sections of this paper.

We now state some basic properties of the curvelet transform. First, note that the curvelet transform works primarily on the high-frequency components of  $f$ , which correspond to fine-scale elements ( $a$  small). The constant factor  $\lambda$ , mentioned above, sets the low-frequency cutoff value, which corresponds to the coarsest scale ( $a = 1$ ). For convenience, here we assume that  $f$  has no low-frequency components below the cutoff value. In practice, when  $f$  has such low-frequency components, the curvelet transform leaves them unchanged, and simply returns them as a residual function  $f_{\text{res}}$ .

Next, we define the reference measure  $d\mu(a, \vec{b}, \vec{\theta}) := \frac{da d\vec{b} d\sigma(\vec{\theta})}{a^{n+1}}$ . This weights the contributions of  $\Gamma_f(a, \vec{b}, \vec{\theta})$  differently according to the scale  $a$ . Intuitively, this is needed because the sectors  $S_{a,\vec{\theta}}$  do not cover the frequency domain uniformly, and the translations of a curvelet  $\gamma_{a,\vec{0},\vec{\theta}}$  (for fixed  $a$  and  $\vec{\theta}$ ) to different locations  $\vec{b}$  do not cover the spatial domain uniformly. Rather,  $(a, \vec{b}, \vec{\theta})$  should be “sampled” in a certain way. To see this, write  $d\mu(a, \vec{b}, \vec{\theta}) = \frac{da}{a} \frac{d\vec{b}}{a^{(n+1)/2}} \frac{d\sigma(\vec{\theta})}{a^{(n-1)/2}}$ . Note that  $\frac{da}{a} = d(\log a)$ , suggesting that we should sample  $\log(a)$  at uniform intervals, i.e., we should set  $a$  equal to powers of 2; we should sample  $\vec{b}$  on a grid in  $\mathbb{R}^n$  whose cells have size  $a \times (\sqrt{a})^{n-1}$ ; and we should sample  $\vec{\theta}$  on a mesh on  $S^{n-1}$  whose cells have size  $(\sqrt{a})^{n-1}$ . Later, when we construct

the discrete curvelet transform, we will use this sampling trick for  $a$  and  $\vec{\theta}$ , in place of the reference measure.

Then we have the following theorems:

**Theorem 1** *Suppose that  $\hat{f}(\vec{k}) = 0$  for all  $|\vec{k}| < 1/\lambda$ . Then we can recover  $f$  from its curvelet transform  $\Gamma_f$ :  $f(\vec{x}) = \int_{a < 1} \Gamma_f(a, \vec{b}, \vec{\theta}) \gamma_{a, \vec{b}, \vec{\theta}}(\vec{x}) d\mu(a, \vec{b}, \vec{\theta})$ .*

**Theorem 2** *Suppose that  $\hat{f}(\vec{k}) = 0$  for all  $|\vec{k}| < 1/\lambda$ . Then the curvelet transform preserves the  $L^2$  norm:  $\int_{\mathbb{R}^n} |f(\vec{x})|^2 d\vec{x} = \int_{a < 1} |\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\mu(a, \vec{b}, \vec{\theta})$ .*

These are straightforward generalizations (to the case of  $\mathbb{R}^n$ ) of results in [8]. We sketch the proofs in Appendix A.

### 3 The Curvelet Transform of a Radial Function

$|\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\mu(a, \vec{b}, \vec{\theta})$  can be interpreted as a probability density over the different scales, locations and directions  $(a, \vec{b}, \vec{\theta})$ . In this section we will develop some tools for understanding where this probability mass is concentrated. We will consider the case where  $f$  has rotational symmetry. Though there is no simple analytic expression for  $\Gamma_f$ , we can deduce certain properties from symmetry, and we can upper-bound the variance of  $\vec{b}$  (this latter point is our main result).

Let  $f$  be a radial function,  $f(\vec{x}) = f_0(|\vec{x}|)$ . Its Fourier transform is also radial,  $\hat{f}(\vec{k}) = F_0(|\vec{k}|)$ , where  $F_0(\rho) = \frac{2\pi}{\rho^{(n-2)/2}} \int_0^\infty J_{(n-2)/2}(2\pi\rho r) f_0(r) r^{n/2} dr$ , and  $J$  is a Bessel function (see, e.g., [29]). We assume that  $f$  is normalized so that  $\int_{\mathbb{R}^n} |f(\vec{x})|^2 d\vec{x} = 1$ .

When  $f$  is radial,  $\Gamma_f$  has the following symmetries:  $\Gamma_f(a, \vec{b}, \vec{\theta}) = \Gamma_f(a, -\vec{b}, -\vec{\theta})$ , and for any rotation  $R$ ,  $\Gamma_f(a, \vec{b}, \vec{\theta}) = \Gamma_f(a, R(\vec{b}), R(\vec{\theta}))$ .

We make a particular choice for the radial and angular windows  $W$  and  $V$ . These windows are  $C^1$  smooth, which is necessary in our analysis of the variance of  $\vec{b}$ . We let  $W(r) = [C_w \sin(\pi \log r)]^2$  if  $1/e \leq r \leq 1$ ; 0 otherwise] and  $V(t) = [C_v \cos(t)]^2$  if  $0 \leq t \leq \pi/2$ ; 0 otherwise], where  $C_w = \sqrt{8/3}$ , and  $C_v = \sqrt{\frac{2(n+2)n}{3S_0}}$ .

**3.1 The probability of observing a scale  $a$ :** First, we claim that the probability of observing a fine-scale element  $a \leq \eta$  is essentially given by the amount of probability mass of  $\hat{f}$  at frequencies above  $1/(\lambda\eta)$ :  $\Pr[a \leq \eta] = \int_{a \leq \eta} |\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\mu(a, \vec{b}, \vec{\theta}) \geq \int_{|\vec{k}| \geq 1/(\lambda\eta)} |\hat{f}(\vec{k})|^2 d\vec{k}$ . This follows from the same argument used to prove Theorem 2. In the case of a radial function, we write this as:

$$\Pr[a \leq \eta] \geq S_0 \int_{1/(\lambda\eta)}^\infty F_0(\rho)^2 \rho^{n-1} d\rho, \quad (2)$$

where  $S_0 = \frac{2\pi^{n/2}}{\Gamma(n/2)}$  is the surface area of the sphere  $S^{n-1} \subset \mathbb{R}^n$ .

**3.2 The location  $\vec{b}$  and direction  $\vec{\theta}$ :** We claim that the location  $\vec{b}$  has expectation value  $\vec{0}$ . To see this, observe that  $E(b_j) = -E(b_j)$ , due to the reflection symmetry of  $\Gamma_f$ ; thus  $E(b_j) = 0$ . Note that this remains true when we condition on the value of  $a$ .

Also, we claim that the direction  $\vec{\theta}$  is uniformly distributed. This follows from the rotational symmetry of  $\Gamma_f$ . Note that this remains true when we condition on the value of  $a$ , and when we condition on the value of  $\vec{b} \cdot \vec{\theta}$  (since this preserves the rotational symmetry).

**3.3 The variance of  $\vec{b}$  perpendicular to  $\vec{\theta}$ :** Finally, we seek to upper-bound the variance of  $\vec{b}$ , in the directions perpendicular to  $\vec{\theta}$ , as well as parallel to  $\vec{\theta}$ . These results are rather complicated, so we defer most of the details to Appendix B.1. However, these results are a basic component of our proofs in Sections 4 and 5, so we will sketch some of the calculations.

First, the variance of  $\vec{b}$  perpendicular to  $\vec{\theta}$  is:

$$\mathbb{E}(\vec{b}^T(I - \vec{\theta}\vec{\theta}^T)\vec{b}) = \int_{S^{n-1}} \int_0^1 \int_{\mathbb{R}^n} (\vec{b}^T(I - \vec{\theta}\vec{\theta}^T)\vec{b}) |\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\vec{b} \frac{da}{a^{n+1}} d\sigma(\vec{\theta}). \quad (3)$$

Note that a similar formula holds when we condition on observing  $a \leq \eta$ .

We can take advantage of rotational symmetry to do the  $\vec{\theta}$  integral. Fix a vector  $\vec{u} = (1, 0, \dots, 0)$ , and for each  $\vec{\theta}$ , let  $R$  be a rotation that maps  $\vec{\theta}$  to  $\vec{u}$ . Then we can replace the expression inside the integral with  $(R(\vec{b})^T(I - \vec{u}\vec{u}^T)R(\vec{b})) |\Gamma_f(a, R(\vec{b}), \vec{u})|^2$ . Then change variables  $\vec{b} \mapsto R^{-1}(\vec{b})$ . The integrand is now independent of  $\vec{\theta}$ , so we can do the  $\vec{\theta}$  integral. We get:

$$\mathbb{E}(\vec{b}^T(I - \vec{\theta}\vec{\theta}^T)\vec{b}) = S_0 \int_0^1 \int_{\mathbb{R}^n} (\vec{b}^T(I - \vec{u}\vec{u}^T)\vec{b}) |\Gamma_f(a, \vec{b}, \vec{u})|^2 d\vec{b} \frac{da}{a^{n+1}}. \quad (4)$$

Now the key idea is to replace integration over the spatial domain with integration over the frequency domain, via Plancherel's theorem. (Recall that curvelets are defined more simply over the frequency domain.) We introduce some new notation,  $\Phi_{a,\vec{\theta}}(\vec{b}) := \Gamma_f(a, \vec{b}, \vec{\theta})$ . By equation (1), the Fourier transform of  $\Phi_{a,\vec{\theta}}$  is given by  $\hat{\Phi}_{a,\vec{\theta}}(\vec{k}) = \hat{f}(\vec{k}) \chi_{a,\vec{\theta}}(\vec{k})$ .

Let  $I_K$  denote the innermost integral in equation (4). Then  $I_K = \sum_{j=2}^n \int_{\mathbb{R}^n} |b_j|^2 |\Phi_{a,\vec{u}}(\vec{b})|^2 d\vec{b}$ . Using Plancherel's theorem, and symmetry with respect to rotations around the  $\vec{u}$  axis, we can write  $I_K = \sum_{j=2}^n \int_{\mathbb{R}^n} |\frac{1}{2\pi i} \frac{\partial}{\partial k_j} \hat{\Phi}_{a,\vec{u}}(\vec{k})|^2 d\vec{k} = \frac{n-1}{(2\pi)^2} \int_{\mathbb{R}^n} |\frac{\partial}{\partial k_2} \hat{\Phi}_{a,\vec{u}}(\vec{k})|^2 d\vec{k}$ .

We can expand out the integral on the right hand side, as follows. Using spherical coordinates  $\vec{k} = (r, \phi_1, \dots, \phi_{n-1})$ , we write  $\hat{\Phi}_{a,\vec{u}}(\vec{k})$  as a product of a radial part and an angular part:  $\hat{\Phi}_{a,\vec{u}}(\vec{k}) = L(r)M(\phi_1)$ , where  $L(r) = F_0(r)W(\lambda ar)$ , and  $M(\phi_1) = V(\phi_1/\sqrt{a})\Lambda_a(\phi_1)$ . Then we have

$$\frac{\partial}{\partial k_2} \hat{\Phi}_{a,\vec{u}}(\vec{k}) = L'(r)M(\phi_1) \frac{\partial r}{\partial k_2} + L(r)M'(\phi_1) \frac{\partial \phi_1}{\partial k_2}, \quad (5)$$

where  $\frac{\partial r}{\partial k_2} = \sin \phi_1 \cos \phi_2$ , and  $\frac{\partial \phi_1}{\partial k_2} = \frac{\cos \phi_1 \cos \phi_2}{r}$ . So we get: (note that  $\hat{\Phi}_{a,\vec{u}}(\vec{k})$  is real)

$$I_K = \frac{n-1}{(2\pi)^2} \int_{S^{n-1}} \int_0^\infty \left( L'(r)M(\phi_1) \sin \phi_1 \cos \phi_2 + L(r)M'(\phi_1)r^{-1} \cos \phi_1 \cos \phi_2 \right)^2 r^{n-1} dr d\sigma(\vec{\phi}). \quad (6)$$

We can then upper-bound these integrals in terms of  $F_0$  (the radial component of  $\hat{f}$ ). See Appendix B.1 for details. The final result is:

$$\begin{aligned} & \mathbb{E}(\vec{b}^T(I - \vec{\theta}\vec{\theta}^T)\vec{b} \mid a \leq \eta) \\ & \leq \frac{1}{\Pr[a \leq \eta]} \frac{n-1}{(2\pi)^2} S_0 \cdot \left[ \frac{1}{2}(n-2) \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-3} dr \right. \\ & \quad + \frac{5}{n-1} \left( \frac{1}{\lambda} \int_{1/(\lambda\eta e)}^\infty F_0'(r)^2 r^{n-2} dr + \frac{17}{\lambda} \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-4} dr \right) \\ & \quad \left. + (2n+9 + \frac{10}{n-3})e\lambda \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-2} dr \right]. \end{aligned} \quad (7)$$

In sections 4 and 5, we will explain how this is used.

**3.4 The variance of  $\vec{b}$  parallel to  $\vec{\theta}$ :** In a similar way, we can upper-bound the variance of  $\vec{b} \cdot \vec{\theta}$ . See Appendix B.2 for details.

## 4 The Ball in $\mathbb{R}^n$

Let  $B$  be a ball in  $\mathbb{R}^n$ , of radius  $\beta$ . In this section we will analyze the curvelet transform of the function  $f(\vec{x}) = [1/\sqrt{\text{vol}(B)} \text{ if } \vec{x} \in B, 0 \text{ otherwise}]$ . This is the wavefunction one gets by quantum-sampling over  $B$ .

We assume  $n \geq 4$ , and we use the window functions  $W$  and  $V$  specified in Section 3. We set the parameter  $\lambda$  to lie in the range  $2\pi\beta e/n \leq \lambda \leq 2 \cdot 2\pi\beta e/n$ . We show the following:

**Theorem 3** *Almost all of the power in  $\hat{f}$  is located at frequencies  $|\vec{k}| \geq 1/\lambda$ :  $\int_{|\vec{k}| \leq 1/\lambda} |\hat{f}(\vec{k})|^2 d\vec{k} < \frac{1}{\pi n}$ . For any  $\eta \leq 1/e^2$ , the probability of observing a fine-scale element  $a \leq \eta$  is lower-bounded by:  $\Pr[a \leq \eta] \geq \frac{e\eta}{14}(1 - \frac{1}{n})$ . Furthermore, if  $\eta \leq (1/2e^2)(1 - \frac{2}{n+2})$ , then the variance of  $\vec{b}$ , in the directions orthogonal / parallel to  $\vec{\theta}$ , conditioned on  $a \leq \eta$ , is upper-bounded by:*

$$E(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) \leq \eta\beta^2(14300 + O(\frac{1}{n})), \quad (8)$$

$$E((\vec{b} \cdot \vec{\theta})^2 \mid a \leq \eta) \leq \beta^2(242 + O(\frac{1}{n})). \quad (9)$$

The first claim shows that only an inverse-polynomial fraction of the probability mass lies below the low-frequency cutoff; this justifies our use of the curvelet transform and Theorems 1 and 2. The second claim shows that, for any sufficiently small constant  $\eta$ , we observe scale  $a \leq \eta$  with constant probability. This is due to the fact that  $f$  has a lot of power at high frequencies (a “heavy tail”), which is caused by the discontinuity of  $f$  along the surface of the ball. (For comparison, one would not observe this behavior if  $f$  were, say, a Gaussian.) The third claim shows that, when  $a \leq \eta$ ,  $\vec{b}$  lies within distance  $O(\sqrt{\eta}\beta)$  of the line that passes through the center of the ball with direction  $\vec{\theta}$ . Also,  $\vec{b}$  lies within distance  $O(\beta)$  of the center (and we expect, though we do not prove, that this distance is also lower bounded by  $\Omega(\beta)$ ).

As mentioned previously, it is remarkable that these bounds do not depend on the dimension  $n$ . It is also interesting that this concentration of probability mass can hold even when the window function  $\chi_{a,\vec{\theta}}$  is only  $C^1$ -smooth. By contrast, in order for  $\Gamma_f$  to be asymptotically rapidly decaying,  $\chi_{a,\vec{\theta}}$  must usually be  $C^k$  or  $C^\infty$ -smooth.

We prove this using our results from Section 3. Note that the curvelet transform behaves in a simple way when we translate the function  $f$ : if  $g(\vec{x}) = f(\vec{x} - \vec{z})$ , then using equation (1), we see that  $\Gamma_g(a, \vec{b}, \vec{\theta}) = \Gamma_f(a, \vec{b} - \vec{z}, \vec{\theta})$ . Thus, without loss of generality, we can assume that the ball  $B$  is centered at the origin. In this case  $f$  is a radial function.

We will be interested in the Fourier transform of  $f$ . We write  $f(\vec{x}) = f_0(|\vec{x}|)$ , where  $f_0(r) = [C \text{ if } r \leq \beta, 0 \text{ otherwise}]$ , and  $C = 1/\sqrt{\text{vol}(B)} = (S_0\beta^n/n)^{-1/2}$ , where  $S_0$  is the surface area of the unit sphere  $S^{n-1}$  in  $\mathbb{R}^n$ . Then the Fourier transform of  $f$  is given by  $\hat{f}(\vec{k}) = F_0(|\vec{k}|)$ , where

$$F_0(\rho) = \sqrt{\frac{n}{S_0}} \frac{1}{\rho^{n/2}} J_{n/2}(2\pi\rho\beta), \quad (10)$$

using the definition from Section 3, and the identity  $\frac{d}{dz}(z^\nu J_\nu(z)) = z^\nu J_{\nu-1}(z)$  (see [1], eqn. (9.1.30)).

The behavior of  $F_0(\rho)$  depends on the behavior of the Bessel function  $J_\nu(z)$  when  $\nu \geq 0$  and  $z \geq 0$ .  $J_\nu(z)$  is very small when  $z \ll \nu$ , it undergoes a transition near  $z \approx \nu$ , and it is approximately



given by  $\sqrt{2/(\pi z)} \cos(z - \frac{1}{2}\nu\pi - \frac{1}{4}\pi)$  when  $z \gg \nu$  (see [1]). Thus, our intuition is that  $F_0(\rho) \approx 0$  when  $\rho \lesssim (n/2)/(2\pi\beta)$ , and  $F_0(\rho) \approx (\text{const})/\rho^{(n+1)/2}$  times an oscillating factor when  $\rho \gtrsim (n/2)/(2\pi\beta)$ .

We now sketch the proof, omitting the details. First we choose  $\lambda \approx (\text{const}) \cdot (2\pi\beta)/n$ , so that very little power lies at frequencies below  $1/\lambda$ . Substituting into (2), we get that  $\Pr[a \leq \eta] \approx (\text{const}) \cdot \eta$ .

Then we bound the variance of  $\vec{b}$  perpendicular to  $\vec{\theta}$ , as follows. (A similar argument holds for the variance of  $\vec{b} \cdot \vec{\theta}$ .) We start with (7), and use the fact that  $r \geq 1/(\lambda\eta e)$  implies  $r \cdot \lambda\eta e \geq 1$ :

$$\begin{aligned} & \mathbb{E}(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) \\ & \leq \frac{1}{\Pr[a \leq \eta]} \frac{n-1}{(2\pi)^2} S_0 \cdot \left[ \frac{1}{2}(n-2)(\lambda\eta e)^2 \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-1} dr \right. \\ & \quad + \frac{5}{n-1} \left( \frac{\lambda\eta e}{\lambda} \int_{1/(\lambda\eta e)}^{\infty} F_0'(r)^2 r^{n-1} dr + \frac{17(\lambda\eta e)^3}{\lambda} \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-1} dr \right) \\ & \quad \left. + (2n+9 + \frac{10}{n-3})e\lambda(\lambda\eta e) \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-1} dr \right]. \end{aligned}$$

A standard calculation shows that  $F_0'(r) = -\sqrt{\frac{n}{S_0}} r^{-n/2} J_{(n/2)+1}(2\pi\beta r)(2\pi\beta)$ , which behaves similarly to  $F_0(r)$ , except that it oscillates differently and is larger by a factor of  $2\pi\beta$ . Thus we can get a reasonable estimate by replacing  $F_0'(r)$  with  $F_0(r)(2\pi\beta)$  in the integral:

$$\begin{aligned} & \mathbb{E}(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) \\ & \lesssim \frac{1}{\Pr[a \leq \eta]} \frac{n-1}{(2\pi)^2} S_0 \cdot \left[ \frac{1}{2}(n-2)(\lambda\eta e)^2 \right. \\ & \quad + \frac{5}{n-1} \left( \frac{\lambda\eta e}{\lambda} (2\pi\beta)^2 + \frac{17(\lambda\eta e)^3}{\lambda} \right) + (2n+9 + \frac{10}{n-3})e\lambda(\lambda\eta e) \Big] \\ & \quad \cdot \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-1} dr \end{aligned}$$

Using the definition of  $\lambda$ ,

$$\begin{aligned} & \mathbb{E}(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) \\ & \lesssim \frac{1}{\Pr[a \leq \eta]} \frac{n-1}{(2\pi)^2} S_0 \cdot \left[ (\text{const}) \frac{(2\pi\beta)^2}{n} \eta^2 \right. \\ & \quad + \left( (\text{const}) \frac{(2\pi\beta)^2}{n} \eta + (\text{const}) \frac{(2\pi\beta)^2}{n^3} \eta^3 \right) + (\text{const}) \frac{(2\pi\beta)^2}{n} \eta \Big] \\ & \quad \cdot \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-1} dr. \end{aligned}$$

Now recall the expression for  $\Pr[a \leq \eta]$  given by (2). We expect the two integrals to roughly cancel out, so we get

$$\mathbb{E}(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) \lesssim O(\beta^2 \eta). \quad (11)$$

This argument can be made rigorous, using known results about Bessel functions  $J_\nu(z)$ . However, there is a technical obstacle: our theorem concerns the case where  $z$  is roughly proportional to  $\nu$ . This is still in the transition regime, and the usual asymptotic expansions for  $J_\nu(z)$  do not work here (they only work when  $z \gtrsim \nu^2$ , or when  $z/\nu$  is some fixed ratio). Fortunately, there are useful bounds on the quantity  $M_\nu(z) := \sqrt{J_\nu(z)^2 + Y_\nu(z)^2}$ , and representations of  $J_\nu(z)$  and  $Y_\nu(z)$  in terms of a modulus and phase, that do work in this regime [1, 30]. This leads to a rigorous proof of our theorem—see the Appendix for details.

## 5 Spherical Shells

We now consider the curvelet transform of a function supported on a thin spherical shell in  $\mathbb{R}^n$ . We will show results similar to the previous section, except that they now depend on the *thickness* of the shell. Intuitively, when the shell is very thin, we can measure very fine-scale elements ( $a$  small) with significant probability, and  $\vec{b}$  is tightly concentrated around the wavefront set.

Without loss of generality, we can assume the shell is centered at the origin (see Section 4). So consider the following function on  $\mathbb{R}^n$ ,  $f(\vec{x}) = [C \text{ if } \beta < |\vec{x}| \leq \beta + \delta, 0 \text{ otherwise}]$ , where  $C = 1/\sqrt{(\beta + \delta)^n B_0 - \beta^n B_0}$ , and  $B_0$  is the volume of the unit ball in  $\mathbb{R}^n$ . This represents a uniform superposition over a spherical shell centered at the origin, with inner radius  $\beta$  and thickness  $\delta$ . We call this a spherical shell with “square” cross-section.

This is the exactly the kind of state that appears in our quantum algorithm. However, it is difficult to analyze, as its Fourier transform involves a linear combination of two Bessel functions oscillating at different rates. We are interested in the case where  $\delta \ll \beta$ . In Appendix D.1, we give a heuristic explanation of why the curvelet transform of this state will be tightly concentrated around the wavefront set. (This holds when  $\delta \lesssim \beta/n$ .)

Here, we give a more rigorous argument, for spherical shells that have “Gaussian” cross-sections—when  $\delta \ll \beta$ , these functions are similar to the above, but they are analytically tractable. We define  $f = C_f g * q$ , where:  $C_f$  is a normalization factor;  $g$  is a Gaussian of width  $\delta$ , that is,  $g(\vec{x}) = \delta^{-n/2} \exp(-\pi|\vec{x}|^2/\delta^2)$ ;  $q$  is the measure supported on the sphere of radius  $\beta$  around the origin, which is obtained by restricting the usual volume measure on  $\mathbb{R}^n$ ; and the star denotes convolution. Intuitively,  $q$  represents a shell with infinitesimal thickness, and  $f$  represents a “smoothed” shell with thickness  $\delta$ .

The Fourier transform of  $f$  is given by  $\hat{f} = C_f \hat{g} \cdot \hat{q}$ , where:  $\hat{g}$  is a Gaussian of width  $1/\delta$ ,  $\hat{g}(\vec{k}) = \delta^{n/2} \exp(-\pi\delta^2|\vec{k}|^2)$ ; and  $\hat{q}$  is given by  $\hat{q}(\vec{k}) = Q_0(|\vec{k}|)$ , where  $Q_0(\rho) = \frac{2\pi}{\rho^{(n/2)-1}} J_{(n/2)-1}(2\pi\beta\rho)\beta^{n/2}$ . Intuitively, the Fourier transform of the spherical shell is somewhat like the Fourier transform of the ball, except that it decays more slowly (i.e., has more power at high frequencies), for frequencies up to roughly  $1/\delta$ ; the power at frequencies above  $1/\delta$  is suppressed by  $\hat{g}$ .

Note that this is quite similar to equation (128), describing a spherical shell with “square” cross-section, when we substitute in the upper and lower bounds on  $C_f$  (to be given later in this section). This suggests that the spherical shell with “square” cross-section can indeed be approximated by one with “Gaussian” cross-section, when  $\delta \lesssim \beta/n$ .

We will prove bounds on the continuous curvelet transform, with the same window functions as in Section 4. We use a slightly different scaling parameter  $\lambda$ : we set  $\lambda = \frac{2\pi\tilde{\beta}e}{n-2}$ , where  $\tilde{\beta}$  is an estimate of the true radius of the shell, which satisfies  $\beta \leq \tilde{\beta} \leq S\beta$ , for some  $S \geq 1$ . We assume that the dimension  $n$  is at least 4, and we assume that the thickness of the shell is small compared to the radius:  $\delta = \varepsilon\beta$ , where  $\varepsilon \leq \min\left(\frac{6}{(n-2)^2}, \frac{1}{n+2}, \frac{1}{en}, \frac{1}{5}\right)$ . (Note that for these values of  $n$ ,  $\frac{1}{en} \leq \frac{1}{n+2}$  and  $\frac{1}{en} \leq \frac{1}{5}$ , so the second and fourth conditions are actually redundant.) Under these assumptions, we prove the following:

**Theorem 4** *Almost all of the power in  $\hat{f}$  is located at frequencies  $|\vec{k}| \geq 1/\lambda$ :  $\int_{|\vec{k}| \leq 1/\lambda} |\hat{f}(\vec{k})|^2 d\vec{k} \leq \frac{\varepsilon}{5}$ . Let  $\eta_c = (\delta/\tilde{\beta})(n-2)/e$ . The probability of observing a fine-scale element  $a \leq \eta_c$  is lower-bounded by:  $\Pr[a \leq \eta_c] > 0.045$ . Furthermore, the variance of  $\vec{b}$ , in the directions orthogonal / parallel to  $\vec{\theta}$ , conditioned on  $a \leq \eta_c$ , is upper-bounded by:*

$$E(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} | a \leq \eta_c) \leq (n-1)\varepsilon\beta^2(507 + O(\frac{1}{n})) \cdot S, \quad (12)$$

$$E((\vec{b} \cdot \vec{\theta})^2 \mid a \leq \eta_c) \leq \beta^2(23 + O(\frac{1}{n^2})). \quad (13)$$

The proof uses a similar strategy to what we showed in Section 4. The intuition is as follows. First write  $\hat{f}(\vec{k}) = F_0(|\vec{k}|)$ . The important difference (compared to Section 4) is that here  $F_0(r)$  decays more slowly, like  $1/r^{(n-1)/2}$ , for  $r \lesssim 1/\delta$ . Substituting into (2), we see that a constant fraction of the probability mass lies at frequencies of order  $1/\delta$ . So with constant probability, we can observe fine-scale elements  $a \leq \eta_c$  where  $\eta_c$  is of order  $\delta/\lambda$ . Note that  $\eta_c \approx \delta/\lambda \approx (\text{const})(\delta/\beta)(n-2) \approx (\text{const})\varepsilon(n-2)$ . So, when the shell is very thin,  $\eta_c$  will be very small, and  $\vec{b}$  will be tightly concentrated around the wavefront set.

The rigorous proof is given in Appendix D.2.

## 6 A Fast Quantum Curvelet Transform

**6.1 The Discrete Curvelet Transform:** First, we describe the discrete curvelet transform, which has been studied in the classical setting [6]. The discrete curvelet transform takes a function  $f(\vec{x})$  and returns a function  $\Gamma_f(a, \vec{b}, \vec{\theta})$ , where both functions are defined over finite domains. This is constructed analogously to the continuous curvelet transform, except that one now uses the discrete Fourier transform on  $(\mathbb{Z}_M)^n$ , and a discrete set of scale/direction pairs  $(a, \vec{\theta})$ .

The discrete Fourier transform is defined as follows. We assume that  $f$  is defined on a domain  $Z$  consisting of a discrete grid in a finite region of  $\mathbb{R}^n$ . For example, let  $Z = (\sigma\mathbb{Z})^n \cap [-L, L]^n$ , the intersection of a tightly-spaced square lattice and a large cube. Also let  $\hat{Z} = (\frac{1}{2L}\mathbb{Z})^n \cap [-\frac{1}{2\sigma}, \frac{1}{2\sigma}]^n$ . The discrete Fourier transform maps  $f$  to a function  $\hat{f}$  defined on  $\hat{Z}$ , as follows:  $\hat{f}(\vec{k}) = (\frac{\sigma}{2L})^{n/2} \sum_{\vec{x} \in Z} f(\vec{x}) e^{-2\pi i \vec{k} \cdot \vec{x}}$ , and  $f(\vec{x}) = (\frac{\sigma}{2L})^{n/2} \sum_{\vec{k} \in \hat{Z}} \hat{f}(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{x}}$ .

One can argue that this approximates the continuous Fourier transform in the following sense. Let  $f_{\text{cont}}$  be a function on  $\mathbb{R}^n$ , and let  $\hat{f}_{\text{cont}}$  be its continuous Fourier transform. Suppose that  $f_{\text{cont}}$  is supported inside the cube  $[-L, L]^n$ , and  $\hat{f}_{\text{cont}}$  has all except an  $\varepsilon$  fraction of its probability mass inside the cube  $[-\frac{1}{2\sigma}, \frac{1}{2\sigma}]^n$ . Then there exists a function  $f_{\text{dis}}$  on  $Z$ , with discrete Fourier transform  $\hat{f}_{\text{dis}}$ , such that  $f_{\text{dis}} \approx \sigma^{n/2} f_{\text{cont}}|_Z$  and  $\hat{f}_{\text{dis}} = (\frac{1}{2L})^{n/2} \hat{f}_{\text{cont}}|_{\hat{Z}}$ , up to errors whose total probability mass is roughly  $\varepsilon$ . See Appendix E.1 for more details.

Now the discrete curvelet transform is given by

$$\Gamma_f(a, \vec{b}, \vec{\theta}) := (\frac{\sigma}{2L})^{n/2} \sum_{\vec{k} \in \hat{Z}} \hat{f}(\vec{k}) \chi_{a, \vec{\theta}}(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{b}}. \quad (14)$$

The “location” variables  $\vec{x}$  and  $\vec{b}$  take values in  $Z$ , and the “scale” and “direction” variables  $(a, \vec{\theta})$  take values in some discrete set  $G$ , which we will describe below. The window functions  $\chi_{a, \vec{\theta}}$  are defined over  $\hat{Z}$ , and are constructed so that  $\sum_{a, \vec{\theta}} \chi_{a, \vec{\theta}}(\vec{k})^2 = 1$ ,  $\forall \vec{k} \in \hat{Z}$ . This ensures that the curvelet transform can be realized as a unitary operation on the space spanned by the states  $|\vec{k}, a, \vec{\theta}\rangle$ .

Recall from Section 2 that each window function  $\chi_{a, \vec{\theta}}$  is supported on a “sector”  $S_{a, \vec{\theta}}$  that has angular width  $\sqrt{a}$ , inner radius  $1/a$  and outer radius  $2/a$ . To satisfy the above condition on  $\chi_{a, \vec{\theta}}$ , we want to choose a discrete set of values  $(a, \vec{\theta})$ , that corresponds to a discrete collection of sectors  $S_{a, \vec{\theta}}$ , that forms a “tiling” of the frequency domain. Intuitively, this is done by setting  $a$  equal to powers of 2, and sampling  $\vec{\theta}$  from a mesh with angular spacing  $\sqrt{a}$  on the sphere  $S_{n-1}$ . Then the sectors  $S_{a, \vec{\theta}}$  fit together nicely, as in Figure 1. (This picture is a slight oversimplification; actually, since we want the window functions  $\chi_{a, \vec{\theta}}$  to decay smoothly to zero, we should make their supports overlap

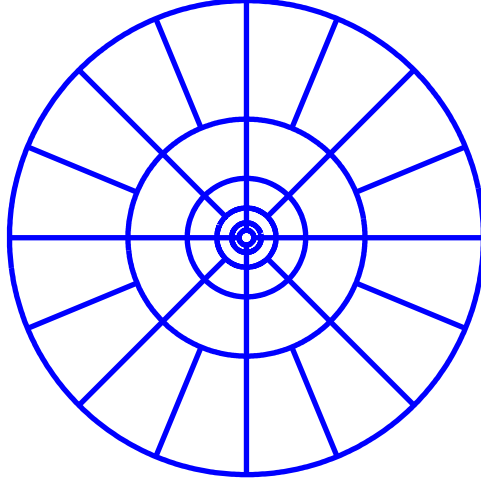


Figure 1: Construction of a 2-D discrete curvelet transform — “tiling” of the frequency domain into sectors  $S_{a,\vec{\theta}}$ , plus special sectors  $S_{low}$  and  $S_{high}$  for very low and very high frequencies. Within the annulus,  $a$  takes on values  $1/2^k$ , for  $k = 0, 1, 2, 3, 4$ . Each sector has inner radius  $2^k = 1/a$ , outer radius  $2^{k+1} = 2/a$ , and angular width  $(\pi/2)/2^{\lfloor k/2 \rfloor} \approx (\pi/2)\sqrt{a}$ .

slightly.) We will describe a construction of this kind in the next section; other constructions were given in [6, 20].

This discretization affects the values of  $a$  and  $\vec{\theta}$ , relative to the continuous case. Intuitively, the “discrete”  $a$  can differ from the “continuous”  $a$  by a constant factor, and the “discrete”  $\vec{\theta}$  can differ from the “continuous”  $\vec{\theta}$  by an additive error of size  $\sqrt{a}$ .

**6.2 The Quantum Curvelet Transform:** The quantum curvelet transform is the unitary operation that maps

$$\sum_{\vec{x}} f(\vec{x}) |\vec{x}\rangle |0, \vec{0}\rangle \mapsto \sum_{a, \vec{b}, \vec{\theta}} \Gamma_f(a, \vec{b}, \vec{\theta}) |\vec{b}\rangle |a, \vec{\theta}\rangle. \quad (15)$$

This can be implemented as follows: first apply the quantum Fourier transform (QFT), then the operation  $\mathcal{X}$  that maps  $|\vec{k}\rangle |0, \vec{0}\rangle \mapsto |\vec{k}\rangle \sum_{a, \vec{\theta}} \chi_{a, \vec{\theta}}(\vec{k}) |a, \vec{\theta}\rangle$ , and then the inverse QFT.

We want to compute this in time polynomial in  $n$  and  $\log(M)$  (where  $M = 2L/\sigma$  is the length of the discrete Fourier transform). This is possible for the QFT. But it is not clear how to perform the operation  $\mathcal{X}$ , for a generic choice of the window functions  $\chi_{a, \vec{\theta}}$ . (Note that we want the functions  $\chi_{a, \vec{\theta}}$  to be  $C^1$ -smooth, so their supports will necessarily overlap; thus the operation  $\mathcal{X}$  must prepare a superposition containing  $2^{\Theta(n)}$  terms.)

Nonetheless, we can perform  $\mathcal{X}$  efficiently in two cases: (1) when the window functions are indicator functions supported on disjoint sets, and (2) when the window functions are smooth “bump” functions that can be expressed as products of 1-D functions using spherical coordinates. The first case has poor analytic properties, but the second case is a reasonable approximation of the curvelets used in Sections 3-5. Thus we get an efficient quantum curvelet transform. See Appendix E.2 for details.

## 7 Quantum Algorithms using the Curvelet Transform

**7.1 Single-shot measurement of a quantum-sample state:** Consider the following problem. Let  $B$  be a ball of (unknown) radius  $\beta$  centered at some (unknown) point  $\vec{c}$  in  $\mathbb{R}^n$ , for  $n \geq 4$ . We are given as input:  $n$ , the dimension;  $\tilde{\beta}$ , an estimate of the radius of the ball (we are promised that  $\tilde{\beta}/2 \leq \beta \leq \tilde{\beta}$ );  $R$ , an outer bound on the location of the center (we are promised that  $|\vec{c}| \leq R$ );  $\mu$ , the desired accuracy of our answer; a description of the set of grid points  $G = (\sigma\mathbb{Z})^n \cap [-L, L]^n$ , such that  $L \geq R + \tilde{\beta}$  and  $\sigma \leq \frac{\pi e}{600} \frac{\mu^2}{\tilde{\beta}} \frac{1}{14300n + Q_1}$  (for some constant  $Q_1$ , to be specified later); and a quantum state  $\frac{1}{\sqrt{|G \cap B|}} \sum_{\vec{x} \in G \cap B} |\vec{x}\rangle$ , that is, a single quantum-sample over the ball  $B$ . We are then asked to output a point  $\vec{z}$  in  $\mathbb{R}^n$ , that lies within distance  $\mu$  of the center  $\vec{c}$ .

We propose the following algorithm. Intuitively, this algorithm uses the curvelet transform to find a line that passes near the center of the ball, then guesses a random point along this line.

**Algorithm 1:**

Let  $|\psi\rangle$  be the input quantum state.

Set  $\eta = \frac{1}{6} \frac{\mu^2}{\tilde{\beta}^2} \frac{1}{14300 + (Q_1/n)}$ , where  $Q_1$  is some constant.

Apply the fast quantum curvelet transform, with  $\lambda = \frac{2\pi\tilde{\beta}e}{n}$ ,  $s_{\min} = 1$ ,  $s_{\max} = \lg \frac{1}{\eta} + 3$ .

Measure the scale  $a = 2^{-s}$ , location  $\vec{b}$  and direction  $\vec{\theta}$ .

If  $a > \eta$ , then return “no answer.”

Set  $Q_3 = \sqrt{3} \sqrt{242 + (Q_2/n)}$ , where  $Q_2$  is some constant.

Guess some  $u \in [-1, 1]$  uniformly at random.

Return the point  $\vec{b}' = \vec{b} + uQ_3\tilde{\beta}\vec{\theta}$ .

We are especially interested in instances where the error  $\mu$  is a constant fraction of the radius  $\beta$ , i.e.,  $\mu = \nu\beta$ , for some fixed  $\nu < 1$ . We conjecture that, for any  $\nu$ , Algorithm 1 solves these instances with probability  $\Omega(\nu^3)$ , independent of the dimension  $n$ . (In other words, the success probability has a “heavy tail.”) This is a sharp contrast to what happens in the classical case: if we choose a single point uniformly at random from the ball, then the success probability is  $\nu^n$ , which is exponentially small in  $n$ . This is because, in high dimensions, most of the volume of the ball lies near its surface. This is bad for classical sampling, but it helps the quantum curvelet transform, which works by finding a line normal to the surface of the ball.

We now show how our results from Section 4 support this conjecture. We prove the following:

**Theorem 5** *Consider a “continuous” analogue of Algorithm 1, using the continuous curvelet transform over  $\mathbb{R}^n$ . This algorithm succeeds with constant probability  $\geq \Omega(\nu^3)$ .*

We will also argue, non-rigorously, that the discrete algorithm will behave like the continuous one, provided that the grid  $G$  is sufficiently fine. When the grid is chosen properly, the discrete algorithm runs in time  $\text{poly}(n, \log R, \log \tilde{\beta}, \log \frac{1}{\mu})$ . See Appendix F.1 for details.

We remark that it should be possible to achieve a better success probability,  $\Omega(\nu^2)$ , using the quantum curvelet transform. Here, we showed that  $\vec{b}$  was within distance  $O(\beta)$  of the center, so in the last step of the algorithm, we simply guessed a point along the line, with success probability  $\Omega(\nu)$ . But in fact,  $\vec{b}$  should lie at distance  $\approx \beta$  from the center, so we should be able to guess one of the two points  $\vec{b} \pm \beta\vec{\theta}$ , with success probability  $\Omega(1)$ .

We also remark that classical sampling becomes more powerful if one is allowed any constant number of samples, instead of just one. By sampling  $k$  random points from the ball and taking

their average, one can find the center with accuracy  $\pm\beta/\sqrt{k}$ , with constant probability. (However, for fixed  $k$ , the success probability does not have a “heavy tail,” i.e., one cannot expect to get better accuracy with significant probability. This is because, in high dimensions, random sampling produces  $k$  vectors that are nearly orthogonal.)

**7.2 Quantum algorithm for finding the center of a radial function:** Let  $f$  be a radial function on  $\mathbb{R}^n$  (where  $n \geq 4$ ), centered at some point  $\vec{c}$ , and taking values in some arbitrary set. Suppose that the level sets of  $f$  are concentric spherical shells of thickness  $\delta$  centered at  $\vec{c}$ , i.e.,  $f$  is constant on each shell, and  $f$  takes on distinct values on different shells. (Note, in previous versions of this paper, we made an additional assumption, that one can efficiently compute the radius of a shell, given the value of  $f$  on that shell. This assumption is no longer needed.)

Consider the following problem. We are given as input:  $n$ , the dimension;  $R$ , an outer bound on the location of the center (we are promised that  $|\vec{c}| \leq R$ );  $\delta$ , the thickness of the spherical shells;  $\mu$ , the desired accuracy of our answer; and an oracle that computes the radial function  $f$ . We are asked to output a point  $\vec{z}$  in  $\mathbb{R}^n$ , that lies within distance  $\mu$  of the center  $\vec{c}$ .

We propose the following algorithm. The basic idea is to prepare a quantum superposition over a large ball around the origin, then measure the value of  $f$  to get a superposition over a spherical shell centered at  $\vec{c}$ , then apply the curvelet transform, and find a line that passes near  $\vec{c}$ . The algorithm does this twice, then returns the point on the first line that lies closest to the second line (note that, with high probability, the two lines are nearly orthogonal).

**Algorithm 2:**

Set  $R' = nR$ . Let  $B$  be a ball of radius  $R'$  around  $\vec{0}$ .

Choose the grid  $G = (\sigma\mathbb{Z})^n \cap [-L, L]^n$ , where  $L = 2R'$  and  $\sigma = \delta/400$ .

For  $i \in \{1, 2\}$ , do the following:

Prepare the state  $\frac{1}{\sqrt{|G \cap B|}} \sum_{\vec{x} \in G \cap B} |\vec{x}\rangle$ , using the methods of [2] or [17].

Compute the value of  $f$  in an auxiliary register, and measure it; call this  $y^{(i)}$ .

Set  $\tilde{\beta}^{(i)} = R' - R$ .

Set  $\eta^{(i)} = (\delta/\tilde{\beta}^{(i)})(n-2)/e$ .

Apply the fast quantum curvelet transform,

with  $\lambda^{(i)} = \frac{2\pi\tilde{\beta}^{(i)}e}{n-2}$ ,  $s_{min}^{(i)} = 1$ ,  $s_{max}^{(i)} = \lg \frac{1}{\eta^{(i)}} + 3$ .

Measure the scale  $a^{(i)} = 2^{-s^{(i)}}$ , location  $\vec{b}^{(i)}$  and direction  $\vec{\theta}^{(i)}$ .

If  $a^{(i)} > \eta^{(i)}$ , then return “no answer.”

End for.

If  $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| > 3/4$ , then return “no answer.”

Set  $r = \vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}$ ,  $s = \vec{\theta}^{(1)} \cdot (\vec{b}^{(1)} - \vec{b}^{(2)})$ ,  $t = \vec{\theta}^{(2)} \cdot (\vec{b}^{(1)} - \vec{b}^{(2)})$ .

Return the point  $\frac{-s+rt}{1-r^2}\vec{\theta}^{(1)} + \vec{b}^{(1)}$ .

We conjecture that Algorithm 2 finds the center with arbitrary precision  $\mu$ , provided that  $\delta$  is sufficiently small, i.e., the radial function  $f$  computed by the oracle is sufficiently “precise.” Let us assume that  $\delta \leq \frac{1}{192} \cdot \frac{\mu^2}{(n-1)^2 R} \cdot \frac{1}{761+(Q_1/n)}$ , for some constant  $Q_1$  to be defined later. We conjecture that Algorithm 2 then finds a solution with constant probability, independent of the dimension  $n$ . Thus, only  $O(1)$  oracle queries are needed. This is an improvement over the classical case, where  $\tilde{\Omega}(n \log \frac{R}{\mu})$  queries are required. (The  $\tilde{\Omega}$  indicates that we are omitting some log factors.)

We now show how our results from Section 5 support this conjecture. We prove the following:

**Theorem 6** *Consider a “continuous” analogue of Algorithm 2, using the continuous curvelet transform over  $\mathbb{R}^n$ . This algorithm succeeds with constant probability.*

We will also argue, non-rigorously, that the discrete algorithm will behave like the continuous one, provided that the grid  $G$  is sufficiently fine. When the grid is chosen properly, the algorithm runs in time  $\text{poly}(n, \log R, \log \frac{1}{\mu})$ . See Appendix F.2 for details.

**7.3 Classical lower bound:** We claim that any classical algorithm for finding the center of a radial function must use at least  $\tilde{\Omega}(n \log \frac{R}{\mu})$  queries. (The  $\tilde{\Omega}$  indicates that we are omitting some log factors.)

Our intuition is as follows. Any algorithm can be described as a decision tree, where each node represents a query to the oracle  $f$ , and the algorithm chooses which branch to follow depending on the oracle's answer. However, the values of the function  $f$  are meaningless by themselves, so when the algorithm receives an answer from the oracle, the algorithm cannot do anything besides comparing this answer with the answers returned previously. Thus, after its  $k$ 'th query, the algorithm can choose one of at most  $k$  distinct branches.

It follows that, if the algorithm makes  $\ell$  queries, the number of possible outputs (i.e., the number of leaves in the tree) is at most  $\ell!$ . In order to solve this problem, however, the algorithm must be able to output at least  $(R/\mu)^n$  different points. So we have  $(R/\mu)^n \leq \ell!$ , which implies  $\ell \geq \tilde{\Omega}(n \log \frac{R}{\mu})$ .

A formal statement and proof of this result is given in Appendix F.3.

**7.4 Finding the center through multiple iterations:** We now describe a variant of Algorithm 2, for finding the center of a radial function. This new algorithm will use multiple iterations, and a larger number of queries, but it has a less demanding requirement on the thickness of the shells that form the level sets of the radial function  $f$ .

First, we describe a single iteration of the new algorithm. We call this procedure  $\text{OneRound}()$ . This is similar to Algorithm 2, but it starts out with a promise that the center lies within distance  $R$  of some point  $\vec{p}$ , and it returns a point  $\vec{q}$  that lies within distance  $\sqrt{R}\sqrt{\mu/2}$  of the center.  $\text{OneRound}()$  also takes a parameter  $S \geq 1$  that controls the accuracy and success probability:  $\text{OneRound}()$  returns a point  $\vec{q}$  (instead of “no answer”) with constant probability, and when this happens, the point  $\vec{q}$  is accurate with probability  $\geq 1 - O(\frac{1}{S})$ .

**Procedure  $\text{OneRound}(R, \vec{p}, S)$ :**

Set  $R' = nSR$ . Let  $B$  be a ball of radius  $R'$  around  $\vec{0}$ .

Choose the grid  $G = (\sigma\mathbb{Z})^n \cap [-L, L]^n$ , where  $L = 2R'$  and  $\sigma = \delta/400$ .

For  $i \in \{1, 2\}$ , do the following:

Prepare the state  $\frac{1}{\sqrt{|G \cap B|}} \sum_{\vec{x} \in G \cap B} |\vec{x}\rangle$ , using the methods of [2] or [17].

Define the function  $g(\vec{x}) = f(\vec{x} + \vec{p})$ .

Compute the value of  $g$  in an auxiliary register, and measure it; call this  $y^{(i)}$ .

Set  $\tilde{\beta}^{(i)} = R' - R$ .

Set  $\eta^{(i)} = (\delta/\tilde{\beta}^{(i)})(n-2)/e$ .

Apply the fast quantum curvelet transform,

with  $\lambda^{(i)} = \frac{2\pi\tilde{\beta}^{(i)}e}{n-2}$ ,  $s_{\min}^{(i)} = 1$ ,  $s_{\max}^{(i)} = \lg \frac{1}{\eta^{(i)}} + 3$ .

Measure the scale  $a^{(i)} = 2^{-s^{(i)}}$ , location  $\vec{b}^{(i)}$  and direction  $\vec{\theta}^{(i)}$ .

If  $a^{(i)} > \eta^{(i)}$ , then return “no answer.”

End for.

If  $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| > 3/4$ , then return “no answer.”

Set  $r = \vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}$ ,  $s = \vec{\theta}^{(1)} \cdot (\vec{b}^{(1)} - \vec{b}^{(2)})$ ,  $t = \vec{\theta}^{(2)} \cdot (\vec{b}^{(1)} - \vec{b}^{(2)})$ .

Return the point  $\vec{q} = \vec{p} + \frac{s+rt}{1-r^2}\vec{\theta}^{(1)} + \vec{b}^{(1)}$ .

Now we describe the full algorithm, with multiple iterations. This algorithm begins with a point at distance  $R$  from the center, then uses  $\text{OneRound}()$  to find a point at distance  $\sqrt{R}\sqrt{\mu/2}$  from the center, and by repeating the procedure, shrinks the distance to  $R^{1/4}(\mu/2)^{3/4}$ ,  $R^{1/8}(\mu/2)^{7/8}$  and so on. It may seem surprising that the distance decreases by more than a constant factor during each iteration. Intuitively, this is because the spherical shells used by the algorithm are not exact dilations of each other. Recall that the shells have different radii  $\beta$ , but they all have the same thickness  $\delta$ . The larger the radius  $\beta$ , the smaller the ratio  $\varepsilon = \delta/\beta$ ; so a larger shell allows a significantly more precise determination of its center. In a sense, the algorithm makes more progress during the early iterations, when the spherical shells are larger.

**Algorithm 3:**

Set  $R_{cur} = R$  and  $\vec{p}_{cur} = \vec{0}$ .

Set  $n_{iter} = \lceil \lg \lg \frac{2R}{\mu} \rceil$ ,  $S = (9.4)n_{iter}$  and  $n_{tries} = 910 \log S$ .

While  $R_{cur} \geq \mu$  do:

    Try running  $\text{OneRound}(R_{cur}, \vec{p}_{cur}, S)$  up to  $n_{tries}$  times.

    If  $\text{OneRound}()$  returns “no answer” on every attempt, then return “no answer.”

    Let  $\vec{q}$  be the point returned by  $\text{OneRound}()$  on one of the successful attempts.

    Set  $R_{cur} = \sqrt{R_{cur}}\sqrt{\mu/2}$  and  $\vec{p}_{cur} = \vec{q}$ .

End while.

Return  $\vec{p}_{cur}$ .

We conjecture that Algorithm 3 will succeed when

$$\delta < \frac{\mu}{128(10 \lg \lg \frac{2R}{\mu})^3 n^2 (507 + \frac{Q_1}{n})}, \quad (16)$$

which is a weaker requirement than that of Algorithm 2, where  $\delta$  had to scale like  $1/R$ . We conjecture that this algorithm then finds a solution with constant probability. Note that this algorithm uses  $O(\lg \lg \frac{2R}{\mu} \lg \lg \lg \frac{2R}{\mu})$  queries, which still beats the classical lower bound of  $\tilde{\Omega}(n \log \frac{R}{\mu})$  queries.

We now show how our results from Section 5 support this conjecture. We prove the following:

**Theorem 7** *Consider a “continuous” analogue of Algorithm 3, using the continuous curvelet transform over  $\mathbb{R}^n$ . This algorithm succeeds with constant probability.*

We will also argue, non-rigorously, that the discrete algorithm will behave like the continuous one, provided that the grid  $G$  is sufficiently fine. When the grid is chosen properly, the algorithm runs in time  $\text{poly}(n, \log R, \log \frac{1}{\mu})$ . See Appendix F.4 for details.

## 8 Conclusions

We introduced the curvelet transform as a tool for quantum algorithms, and demonstrated how it can solve problems involving geometric objects in  $\mathbb{R}^n$ . We showed that: (1) for functions with radial symmetry, the continuous curvelet transform concentrates probability mass near the wavefront set; (2) a quantum curvelet transform (which is a discrete approximation of the continuous curvelet transform) can be implemented efficiently; (3) this leads to quantum algorithms for approximately finding the center of a ball in  $\mathbb{R}^n$ , given a single quantum-sample state, and for exactly finding the center of a radial function in  $\mathbb{R}^n$ , using  $O(1)$  oracle queries.



There are several ways in which these results might be extended. Perhaps one can adapt these quantum algorithms to solve more general problems, like finding the center of an ellipsoid. Perhaps the quantum speed-up can be amplified using a recursive construction, as in [4, 18].

A general open problem is to understand the behavior of the curvelet transform on more complicated shapes. Can one prove that the probability mass of the curvelet transform is concentrated near the wavefront set, for arbitrary functions on  $\mathbb{R}^n$ ? That would generalize the results of this paper, [10] and [7]. Also, can one rigorously bound the approximation of the continuous curvelet transform by a discrete one?

Another problem is to find new quantum algorithms based on the curvelet transform. For example, can one construct a curvelet transform over  $\mathbb{F}_q^n$ , that could help to solve hidden polynomial problems [12]? Are there quantum states with “wavefront” features, from which the curvelet transform could extract useful information? Some candidates are quantum-sample states over convex polytopes [2, 17], and states produced by the evolution of a quantum walk.

One might also try to use the output of the curvelet transform in a more sophisticated way. In this paper, we simply measured the output state, and we made very little use of the scale variable  $a$ , which measures the “sharpness” of the wavefront discontinuity.

*Acknowledgements:* The author is grateful to R. Koenig, J. Preskill, L. Schulman, A. Childs, D. Meyer, N. Wallach, S. Jordan, E. Candès, U. Vazirani (who suggested the iterative algorithm in section 7.4), Z. Landau, D. Aharonov (who pointed out reference [4]), E. Eban, T. Vidick, and the anonymous referees, for helpful discussions and comments. Supported by an NSF Mathematical Sciences Postdoctoral Fellowship.

## References

- [1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. U.S. National Bureau of Standards, 1972 (tenth printing).
- [2] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *STOC*, pages 20–29, 2003.
- [3] D. Bacon, A. M. Childs, and W. van Dam. From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups. In *FOCS*, pages 469–478, 2005.
- [4] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM J. Comput.*, 26(5):1411–1473, 1997.
- [5] F.J. Blanco-Silva. A generalized curvelet transform. Approximation properties. Univ. of S. Carolina, technical report, 2008.
- [6] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Model. Simul.*, 5:861–899, 2005.
- [7] E. J. Candès and D. L. Donoho. Continuous curvelet transform: I. resolution of the wavefront set. *Appl. Comput. Harmon. Anal.*, 19:162–197, 2003.
- [8] E. J. Candès and D. L. Donoho. Continuous curvelet transform: II. discretization and frames. *Appl. Comput. Harmon. Anal.*, 19:198–222, 2003.

- [9] E.J. Candès and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Comm. Pure Appl. Math.*, 58:1472–1528, 2004.
- [10] E.J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise-c2 singularities. *Comm. Pure Appl. Math.*, 57:219–266, 2002.
- [11] A. Chefles. Quantum state discrimination. *Contemporary Physics*, 41(6):401–424, 2000.
- [12] A. M. Childs, L. J. Schulman, and U. V. Vazirani. Quantum algorithms for hidden nonlinear structures. In *FOCS*, pages 395–404, 2007.
- [13] A.M. Childs and W. van Dam. Quantum algorithms for algebraic problems. ArXiv preprint 0812.0380, to appear in *Rev. Mod. Phys.*, 2008.
- [14] T. Decker, J. Draisma, and P. Wocjan. Efficient quantum algorithm for identifying hidden polynomials. arXiv:0706.1219v3 [quant-ph], 2007.
- [15] A. Fijany and C.P. Williams. Quantum wavelet transforms: Fast algorithms and complete circuits. arXiv:quant-ph/9809004v1, 1998.
- [16] M. H. Freedman. Poly-locality in quantum computing. arXiv:quant-ph/0001077, 2000.
- [17] L. Grover and T. Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. arXiv:quant-ph/0208112v1, 2002.
- [18] S. Hallgren and A. Harrow. Superpolynomial speedups based on almost any quantum circuit. In *ICALP*, pages 782–795, 2008.
- [19] S. P. Jordan. Fast quantum algorithm for numerical gradient estimation. *Phys. Rev. Lett.*, 95(5):050501, 2005.
- [20] L. Demanet L. Ying and E. Candès. 3d discrete curvelet transform. In *Proc. SPIE Wavelets XI*, 2005.
- [21] Y.-K. Liu. Quantum algorithms using the curvelet transform. ArXiv preprint 0810.4968, 2008.
- [22] A. Montanaro. Quantum algorithms for shifted subset problems. arXiv:0806.3362v2 [quant-ph], 2008.
- [23] C. Moore, A. Russell, and P. Sniady. On the impossibility of a quantum sieve algorithm for graph isomorphism. In *STOC*, pages 536–545, 2007.
- [24] L. Rademacher and S. Vempala. Dispersion of mass and the complexity of randomized geometric algorithms. In *FOCS*, pages 729–738, 2006.
- [25] O. Regev. Quantum computation and lattice problems. *SIAM J. Comput.*, 33(3):738–760, 2004.
- [26] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997.
- [27] H. F. Smith. A hardy space for fourier integral operators. *J. Geom. Anal.*, 8(4):629–653, 1998.
- [28] E. M. Stein and T. S. Murphy. *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.

- [29] E.M. Stein and R. Shakarchi. *Fourier Analysis: An Introduction*. Princeton University Press, 2003.
- [30] G.N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge, 1962 (2nd ed.).
- [31] P. Høyer. Efficient quantum transforms. arXiv:quant-ph/9702028v1, 1997.
- [32] P. Høyer and J. Neerbek. Bounds on quantum ordered searching. arXiv:quant-ph/0009032v2, 2000.

## A The Continuous Curvelet Transform

We now sketch the proofs of Theorems 1 and 2.

First, for any  $a$  and  $\vec{\theta}$ , let us define

$$g_{a,\vec{\theta}}(\vec{x}) := \int_{\mathbb{R}^n} \langle \gamma_{a,\vec{b},\vec{\theta}}, f \rangle \gamma_{a,\vec{b},\vec{\theta}}(\vec{x}) d\vec{b}. \quad (17)$$

We claim that

$$\hat{g}_{a,\vec{\theta}}(\vec{k}) = |\chi_{a,\vec{\theta}}(\vec{k})|^2 \hat{f}(\vec{k}). \quad (18)$$

To see this, write

$$\langle \gamma_{a,\vec{b},\vec{\theta}}, f \rangle = \int_{\mathbb{R}^n} \gamma_{a,\vec{0},\vec{\theta}}^*(\vec{y} - \vec{b}) f(\vec{y}) d\vec{y} = (\tilde{\gamma}_{a,\vec{0},\vec{\theta}}^* * f)(\vec{b}),$$

where we define

$$\tilde{\gamma}_{a,\vec{0},\vec{\theta}}(\vec{x}) = \gamma_{a,\vec{0},\vec{\theta}}(-\vec{x}).$$

Thus we can write  $g_{a,\vec{\theta}}(\vec{x})$  as

$$g_{a,\vec{\theta}}(\vec{x}) = \int_{\mathbb{R}^n} \gamma_{a,\vec{0},\vec{\theta}}(\vec{x} - \vec{b}) (\tilde{\gamma}_{a,\vec{0},\vec{\theta}}^* * f)(\vec{b}) d\vec{b} = (\gamma_{a,\vec{0},\vec{\theta}} * \tilde{\gamma}_{a,\vec{0},\vec{\theta}}^* * f)(\vec{x}).$$

Taking the Fourier transform, we get

$$\hat{g}_{a,\vec{\theta}}(\vec{k}) = \hat{\gamma}_{a,\vec{0},\vec{\theta}}(\vec{k}) (\tilde{\gamma}_{a,\vec{0},\vec{\theta}}^*)^\wedge(\vec{k}) \hat{f}(\vec{k}) = |\hat{\gamma}_{a,\vec{0},\vec{\theta}}(\vec{k})|^2 \hat{f}(\vec{k}),$$

using the fact that

$$(\tilde{\gamma}_{a,\vec{0},\vec{\theta}}^*)^\wedge(\vec{k}) = \int_{\mathbb{R}^n} \gamma_{a,\vec{0},\vec{\theta}}^*(-\vec{x}) e^{-2\pi i \vec{k} \cdot \vec{x}} d\vec{x} = \int_{\mathbb{R}^n} \gamma_{a,\vec{0},\vec{\theta}}^*(\vec{x}) e^{2\pi i \vec{k} \cdot \vec{x}} d\vec{x} = (\hat{\gamma}_{a,\vec{0},\vec{\theta}}(\vec{k}))^*.$$

This proves equation (18).

Next, we claim that, for all  $|\vec{k}| \geq 1/\lambda$ ,

$$\int_0^1 \int_{S^{n-1}} |\chi_{a,\vec{\theta}}(\vec{k})|^2 d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} = 1. \quad (19)$$

To see this, proceed as follows. Fix  $\vec{k}$ , and write  $\vec{\theta}$  in spherical coordinates centered around  $\vec{k}$ , such that  $\theta_1$  is the angle between  $\vec{\theta}$  and  $\vec{k}$ . Then we have

$$\chi_{a,\vec{\theta}}(\vec{k}) = W(\lambda a |\vec{k}|) V(\theta_1 / \sqrt{a}) a^{(n+1)/4} \left( \frac{\sin(\theta_1 / \sqrt{a}) \sqrt{a}}{\sin(\theta_1)} \right)^{(n-2)/2}.$$

Then substitute into the integral:

$$\begin{aligned}
& \int_0^1 \int_{S^{n-1}} |\chi_{a,\vec{\theta}}(\vec{k})|^2 d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} \\
&= \int_0^1 \int_{S^{n-1}} W(\lambda a |\vec{k}|)^2 V(\theta_1/\sqrt{a})^2 a^{(n+1)/2} \left( \frac{\sin(\theta_1/\sqrt{a})}{\sin(\theta_1)} \right)^{n-2} a^{(n-2)/2} d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} \\
&= \int_0^1 \int_{S^{n-1}} W(\lambda a |\vec{k}|)^2 V(\theta_1/\sqrt{a})^2 \sin^{n-2}(\theta_1/\sqrt{a}) d\theta_1 d\sigma(\theta_2, \dots, \theta_{n-1}) \frac{da}{a^{3/2}}.
\end{aligned}$$

Note that  $V(\theta_1/\sqrt{a})$  is nonzero only when  $\theta_1 \in [0, \pi\sqrt{a}]$ , so we can restrict the integral to this range. Then change variables,  $\theta'_1 = \theta_1/\sqrt{a}$ , to get:

$$\begin{aligned}
& \int_0^1 \int_{S^{n-1}} |\chi_{a,\vec{\theta}}(\vec{k})|^2 d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} \\
&= \int_0^1 \int_{S^{n-1}} W(\lambda a |\vec{k}|)^2 V(\theta'_1)^2 \sin^{n-2}(\theta'_1) d\theta'_1 \sqrt{a} d\sigma(\theta_2, \dots, \theta_{n-1}) \frac{da}{a^{3/2}} \\
&= \int_0^1 \int_{S^{n-1}} W(\lambda a |\vec{k}|)^2 V(\theta_1)^2 d\sigma(\vec{\theta}) \frac{da}{a} \\
&= \int_0^{\lambda|\vec{k}|} W(a)^2 \frac{da}{a} \int_{S^{n-1}} V(\theta_1)^2 d\sigma(\vec{\theta}) \\
&= 1,
\end{aligned}$$

using the admissibility conditions. This proves equation (19).

We now prove Theorem 1. We write

$$\begin{aligned}
& \int_{a < 1} \Gamma_f(a, \vec{b}, \vec{\theta}) \gamma_{a,\vec{b},\vec{\theta}}(\vec{x}) d\mu(a, \vec{b}, \vec{\theta}) \\
&= \int_0^1 \int_{S^{n-1}} \int_{\mathbb{R}^n} \langle \gamma_{a,\vec{b},\vec{\theta}}, f \rangle \gamma_{a,\vec{b},\vec{\theta}}(\vec{x}) d\vec{b} d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} \\
&= \int_0^1 \int_{S^{n-1}} g_{a,\vec{\theta}}(\vec{x}) d\sigma(\vec{\theta}) \frac{da}{a^{n+1}},
\end{aligned}$$

and we claim that this equals  $f(\vec{x})$ . Taking the Fourier transform, we get

$$\int_0^1 \int_{S^{n-1}} \hat{g}_{a,\vec{\theta}}(\vec{k}) d\sigma(\vec{\theta}) \frac{da}{a^{n+1}},$$

and we claim that this equals  $\hat{f}(\vec{k})$ . Using equations (18) and (19), we rewrite this integral as:

$$\int_0^1 \int_{S^{n-1}} |\chi_{a,\vec{\theta}}(\vec{k})|^2 \hat{f}(\vec{k}) d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} = \hat{f}(\vec{k}).$$

(The last equality holds because of (19) when  $|\vec{k}| \geq 1/\lambda$ , and because  $\hat{f}(\vec{k}) = 0$  when  $|\vec{k}| < 1/\lambda$ .) This proves Theorem 1.

Finally, we prove Theorem 2. We write

$$\int_{a < 1} |\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\mu(a, \vec{b}, \vec{\theta}) = \int_0^1 \int_{S^{n-1}} \int_{\mathbb{R}^n} |\langle \gamma_{a,\vec{b},\vec{\theta}}, f \rangle|^2 d\vec{b} d\sigma(\vec{\theta}) \frac{da}{a^{n+1}}.$$

We rewrite the innermost integral, applying some of the identities used to prove (18):

$$\begin{aligned}
\int_{\mathbb{R}^n} |\langle \gamma_{a, \vec{b}, \vec{\theta}}, f \rangle|^2 d\vec{b} &= \int_{\mathbb{R}^n} |(\hat{\gamma}_{a, \vec{0}, \vec{\theta}}^* f)(\vec{b})|^2 d\vec{b} \\
&= \int_{\mathbb{R}^n} |(\hat{\gamma}_{a, \vec{0}, \vec{\theta}}^* \hat{f})(\vec{k})|^2 d\vec{k} \\
&= \int_{\mathbb{R}^n} |(\hat{\gamma}_{a, \vec{0}, \vec{\theta}}(\vec{k}))^* \hat{f}(\vec{k})|^2 d\vec{k} \\
&= \int_{\mathbb{R}^n} |\chi_{a, \vec{\theta}}(\vec{k})|^2 |\hat{f}(\vec{k})|^2 d\vec{k}.
\end{aligned}$$

Substituting in, and using (19), we get:

$$\begin{aligned}
&\int_{a < 1} |\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\mu(a, \vec{b}, \vec{\theta}) \\
&= \int_0^1 \int_{S^{n-1}} \int_{\mathbb{R}^n} |\chi_{a, \vec{\theta}}(\vec{k})|^2 |\hat{f}(\vec{k})|^2 d\vec{k} d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} \\
&= \int_{\mathbb{R}^n} \int_0^1 \int_{S^{n-1}} |\chi_{a, \vec{\theta}}(\vec{k})|^2 d\sigma(\vec{\theta}) \frac{da}{a^{n+1}} |\hat{f}(\vec{k})|^2 d\vec{k} \\
&= \int_{\mathbb{R}^n} |\hat{f}(\vec{k})|^2 d\vec{k} = \int_{\mathbb{R}^n} |f(\vec{x})|^2 d\vec{x}.
\end{aligned}$$

This proves Theorem 2.

## B The Curvelet Transform of a Radial Function

### B.1 The variance of $\vec{b}$ perpendicular to $\vec{\theta}$

This is a continuation of Section 3.3. Recall that we have  $L(r) = F_0(r)W(\lambda ar)$ , and  $M(\phi_1) = V(\phi_1/\sqrt{a})\Lambda_a(\phi_1)$ . Then we write:

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial k_2} \hat{\Phi}_{a, \vec{u}}(\vec{k}) \right|^2 d\vec{k} \\
&= \int_{S^{n-1}} \int_0^\infty \left( L'(r)M(\phi_1) \sin \phi_1 \cos \phi_2 + \right. \\
&\quad \left. L(r)M'(\phi_1)r^{-1} \cos \phi_1 \cos \phi_2 \right)^2 r^{n-1} dr d\sigma(\vec{\phi}) \\
&= I_{Ar}I_{A1}I_2 + 2I_{Br}I_{B1}I_2 + I_{Cr}I_{C1}I_2,
\end{aligned} \tag{20}$$

where we define

$$I_{Ar} = \int_0^\infty L'(r)^2 r^{n-1} dr, \quad I_{A1} = \int_0^\pi M(\phi_1)^2 \sin^n \phi_1 d\phi_1, \tag{21}$$

$$I_{Br} = \int_0^\infty L'(r)L(r)r^{n-2} dr, \quad I_{B1} = \int_0^\pi M'(\phi_1)M(\phi_1) \cos \phi_1 \sin^{n-1} \phi_1 d\phi_1, \tag{22}$$

$$I_{Cr} = \int_0^\infty L(r)^2 r^{n-3} dr, \quad I_{C1} = \int_0^\pi M'(\phi_1)^2 \cos^2 \phi_1 \sin^{n-2} \phi_1 d\phi_1, \tag{23}$$

$$I_2 = \int_{S^{n-2}} \cos^2 \phi_2 d\sigma(\phi_2, \dots, \phi_{n-1}). \tag{24}$$

This shows that the variance of  $\vec{b}$  perpendicular to  $\vec{\theta}$  is:

$$\mathbb{E}(\vec{b}^T(I - \vec{\theta}\vec{\theta}^T)\vec{b}) = \frac{n-1}{(2\pi)^2} S_0 \int_0^1 \left( I_{Ar} I_{A1} I_2 + 2I_{Br} I_{B1} I_2 + I_{Cr} I_{C1} I_2 \right) \frac{da}{a^{n+1}}. \quad (25)$$

A similar formula gives the variance conditioned on observing  $a \leq \eta$ :

$$\mathbb{E}(\vec{b}^T(I - \vec{\theta}\vec{\theta}^T)\vec{b} \mid a \leq \eta) = \frac{1}{\Pr[a \leq \eta]} \frac{n-1}{(2\pi)^2} S_0 \int_0^\eta \left( I_{Ar} I_{A1} I_2 + 2I_{Br} I_{B1} I_2 + I_{Cr} I_{C1} I_2 \right) \frac{da}{a^{n+1}}. \quad (26)$$

We will now upper-bound the various integrals appearing on the right hand side of equation (26).

### B.1.1

We begin with the integral  $\int_0^\eta I_{Br} I_{B1} I_2 da / a^{n+1}$ . A straightforward calculation shows that

$$I_2 = \frac{1}{n-1} \int_{S^{n-2}} d\sigma(\phi_2, \dots, \phi_{n-1}). \quad (27)$$

We can simplify  $I_{B1}$ , by integrating by parts:  $I_{B1} = - \int_0^\pi \frac{1}{2} M(\phi_1)^2 (-1 + n \cos^2 \phi_1) \sin^{n-2} \phi_1 d\phi_1$ . Substituting in the definition of  $M(\phi_1)$ , changing variables, and using the fact that  $V$  is supported on the interval  $[0, \pi/2]$ , we get:

$$\begin{aligned} I_{B1} &= - \int_0^\pi \frac{1}{2} V(\phi_1/\sqrt{a})^2 a^{(2n-1)/2} \left( \frac{\sin(\phi_1/\sqrt{a})}{\sin \phi_1} \right)^{n-2} (-1 + n \cos^2 \phi_1) \sin^{n-2} \phi_1 d\phi_1 \\ &= - \int_0^\pi \frac{1}{2} V(\phi_1/\sqrt{a})^2 a^{(2n-1)/2} \sin(\phi_1/\sqrt{a})^{n-2} (-1 + n \cos^2 \phi_1) d\phi_1 \\ &= - \int_0^{\pi/\sqrt{a}} \frac{1}{2} V(\omega_1)^2 a^n \sin^{n-2} \omega_1 (-1 + n \cos^2(\omega_1 \sqrt{a})) d\omega_1 \\ &= \int_0^{\pi/2} \frac{1}{2} V(\omega_1)^2 a^n \sin^{n-2} \omega_1 (1 - n \cos^2(\omega_1 \sqrt{a})) d\omega_1, \end{aligned}$$

hence

$$|I_{B1}| \leq \frac{1}{2} a^n (n-1) \int_0^{\pi/2} V(\omega_1)^2 \sin^{n-2} \omega_1 d\omega_1.$$

Combining this with  $I_2$ , we get:

$$|I_{B1} I_2| \leq \frac{1}{2} a^n \int_{S^{n-1}} V(\omega_1)^2 d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) = \frac{1}{2} a^n. \quad (28)$$

Next, we can simplify  $I_{Br}$ , by integrating by parts:  $I_{Br} = - \int_0^\infty \frac{1}{2} L(r)^2 (n-2) r^{n-3} dr$ . Combining this with  $I_{B1}$  and  $I_2$ , substituting in the definition of  $L(r)$ , and exchanging the integrals, we get:

$$\begin{aligned} \left| \int_0^\eta I_{Br} I_{B1} I_2 \frac{da}{a^{n+1}} \right| &\leq \int_0^\eta |I_{Br}| \frac{1}{2} \frac{da}{a} = \int_0^\eta \int_0^\infty \frac{1}{2} L(r)^2 (n-2) r^{n-3} dr \frac{1}{2} \frac{da}{a} \\ &= \int_0^\eta \int_0^\infty \frac{1}{2} F_0(r)^2 W(\lambda ar)^2 (n-2) r^{n-3} dr \frac{1}{2} \frac{da}{a} \\ &\leq \int_0^\infty \frac{1}{2} F_0(r)^2 \int_0^\eta W(\lambda ar)^2 \frac{1}{2} \frac{da}{a} (n-2) r^{n-3} dr. \end{aligned}$$

By the definition of  $W$ ,

$$\int_0^\eta W(\lambda ar)^2 \frac{1}{2} \frac{da}{a} = \int_0^{\lambda \eta r} W(\alpha)^2 \frac{1}{2} \frac{d\alpha}{\alpha} \leq \frac{1}{2}, \quad (29)$$

and vanishes when  $r \leq 1/(\lambda \eta e)$ . Thus we have:

$$\left| \int_0^\eta I_{Br} I_{B1} I_2 \frac{da}{a^{n+1}} \right| \leq \frac{1}{4} (n-2) \int_{1/(\lambda \eta e)}^\infty F_0(r)^2 r^{n-3} dr. \quad (30)$$

### B.1.2

Next, consider the integral  $\int_0^\eta I_{Ar} I_{A1} I_2 da / a^{n+1}$ . We already have a bound for  $I_2$ . For  $I_{A1}$ , we write:

$$\begin{aligned} I_{A1} &= \int_0^\pi V(\phi_1 / \sqrt{a})^2 a^{(2n-1)/2} \left( \frac{\sin(\phi_1 / \sqrt{a})}{\sin \phi_1} \right)^{n-2} \sin^n \phi_1 d\phi_1 \\ &= \int_0^{\pi/\sqrt{a}} V(\omega_1)^2 a^n \sin^{n-2} \omega_1 \sin^2(\omega_1 \sqrt{a}) d\omega_1. \end{aligned}$$

Using the fact that  $V$  is supported on  $[0, \pi/2]$ , and the simple bound  $\sin^2(\omega_1 \sqrt{a}) \leq \omega_1^2 a \leq \frac{\pi^2}{4} a$ , we get:

$$0 \leq I_{A1} \leq \int_0^{\pi/2} V(\omega_1)^2 a^n \sin^{n-2} \omega_1 d\omega_1 \frac{\pi^2}{4} a.$$

Combining with  $I_2$ , we get:

$$0 \leq I_{A1} I_2 \leq \frac{\pi^2}{4} \frac{a^{n+1}}{n-1} \int_{S^{n-1}} V(\omega_1)^2 d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) = \frac{\pi^2}{4} \frac{a^{n+1}}{n-1}. \quad (31)$$

We now turn to  $I_{Ar}$ . First, combining with  $I_{A1}$  and  $I_2$ , we have

$$0 \leq \int_0^\eta I_{Ar} I_{A1} I_2 \frac{da}{a^{n+1}} \leq \int_0^\eta I_{Ar} da \cdot \frac{\pi^2}{4(n-1)}. \quad (32)$$

We can upper-bound  $I_{Ar}$  as follows. Note that, for any two  $L^2$  functions, the Cauchy-Schwarz inequality implies that

$$\|f + g\|^2 = \|f\|^2 + 2\langle f, g \rangle + \|g\|^2 \leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 \leq 2\|f\|^2 + 2\|g\|^2; \quad (33)$$

in the last step we used the arithmetic-geometric mean inequality,  $\sqrt{ab} \leq \frac{a+b}{2}$  for  $a, b \geq 0$ , with  $a = \|f\|^2$  and  $b = \|g\|^2$ . Thus we can write

$$\begin{aligned} 0 \leq I_{Ar} &= \int_0^\infty \left( F_0'(r) W(\lambda ar) + F_0(r) W'(\lambda ar) \lambda a \right)^2 r^{n-1} dr \\ &= \|G_1 + G_2\|^2 \leq 2\|G_1\|^2 + 2\|G_2\|^2, \end{aligned} \quad (34)$$

where we define

$$G_1(r) = F_0'(r) W(\lambda ar) r^{(n-1)/2} \quad (35)$$

$$G_2(r) = F_0(r) W'(\lambda ar) \lambda ar^{(n-1)/2}. \quad (36)$$

Thus we have

$$0 \leq \int_0^\eta I_{Ar} I_{A1} I_2 \frac{da}{a^{n+1}} \leq \left( \int_0^\eta \|G_1\|^2 da + \int_0^\eta \|G_2\|^2 da \right) \cdot \frac{\pi^2}{2(n-1)}. \quad (37)$$

We then want to upper-bound the integrals  $\int_0^\eta \|G_1\|^2 da$  and  $\int_0^\eta \|G_2\|^2 da$ .

For the first one, we have:

$$\begin{aligned} \int_0^\eta \|G_1\|^2 da &= \int_0^\eta \int_0^\infty F_0'(r)^2 W(\lambda ar)^2 r^{n-1} dr da \\ &= \int_0^\infty F_0'(r)^2 \int_0^\eta W(\lambda ar)^2 da r^{n-1} dr. \end{aligned}$$

Using the fact that  $W$  is supported on  $[1/e, 1]$ , we can write

$$\int_0^\eta W(\lambda ar)^2 da = \int_0^{\lambda \eta r} W(\alpha)^2 d\alpha \frac{1}{\lambda r} \leq \int_0^{\lambda \eta r} W(\alpha)^2 \frac{d\alpha}{\alpha} \frac{1}{\lambda r} \leq \frac{1}{\lambda r}, \quad (38)$$

and vanishes when  $r \leq 1/(\lambda \eta e)$ . Hence,

$$\int_0^\eta \|G_1\|^2 da \leq \frac{1}{\lambda} \int_{1/(\lambda \eta e)}^\infty F_0'(r)^2 r^{n-2} dr. \quad (39)$$

For the second integral, we have:

$$\begin{aligned} \int_0^\eta \|G_2\|^2 da &= \int_0^\eta \int_0^\infty F_0(r)^2 W'(\lambda ar)^2 \lambda^2 a^2 r^{n-1} dr da \\ &= \int_0^\infty F_0(r)^2 \int_0^\eta W'(\lambda ar)^2 \lambda^2 a^2 da r^{n-1} dr. \end{aligned}$$

Note that the derivative of  $W$  is given by

$$W'(r) = \begin{cases} C_w \sin(2\pi \log r) \pi / r, & 1/e \leq r \leq 1, \\ 0, & \text{otherwise,} \end{cases} \quad (40)$$

where  $C_w = \sqrt{8/3}$ . So we can write

$$\int_0^\eta W'(\lambda ar)^2 \lambda^2 a^2 da = \int_0^{\lambda \eta r} W'(\alpha)^2 \alpha^2 d\alpha \frac{1}{\lambda r^3} \leq \int_{1/e}^1 C_w^2 (\pi/\alpha)^2 \alpha^2 d\alpha \frac{1}{\lambda r^3} = \frac{8}{3} \pi^2 (1 - \frac{1}{e}) \frac{1}{\lambda r^3}, \quad (41)$$

and vanishes when  $r \leq 1/(\lambda \eta e)$ . Hence,

$$\int_0^\eta \|G_2\|^2 da \leq \frac{8}{3} \pi^2 (1 - \frac{1}{e}) \frac{1}{\lambda} \int_{1/(\lambda \eta e)}^\infty F_0(r)^2 r^{n-4} dr. \quad (42)$$



### B.1.3

Finally, we consider the integral  $\int_0^\eta I_{C1} I_{C1} I_2 da/a^{n+1}$ . We already have a bound for  $I_2$ . For  $I_{C1}$  we can write:

$$\begin{aligned} 0 \leq I_{C1} &= \int_0^\pi \left( V'(\phi_1/\sqrt{a})(1/\sqrt{a})\Lambda_a(\phi_1) + V(\phi_1/\sqrt{a})\Lambda'_a(\phi_1) \right)^2 \cos^2 \phi_1 \sin^{n-2} \phi_1 d\phi_1 \\ &= \|U_1 + U_2\|^2 \leq 2\|U_1\|^2 + 2\|U_2\|^2, \end{aligned} \quad (43)$$

where we define

$$U_1(\phi_1) = V'(\phi_1/\sqrt{a})(1/\sqrt{a})\Lambda_a(\phi_1) \cos \phi_1 \sin^{(n-2)/2} \phi_1 \quad (44)$$

$$U_2(\phi_1) = V(\phi_1/\sqrt{a})\Lambda'_a(\phi_1) \cos \phi_1 \sin^{(n-2)/2} \phi_1. \quad (45)$$

Then

$$0 \leq I_{C1} I_2 \leq 2\|U_1\|^2 I_2 + 2\|U_2\|^2 I_2. \quad (46)$$

We now evaluate  $\|U_1\|^2$  and  $\|U_2\|^2$ .

For  $\|U_1\|^2$ , we can write

$$\begin{aligned} \|U_1\|^2 &= \int_0^\pi V'(\phi_1/\sqrt{a})^2 a^{-1} a^{(2n-1)/2} \left( \frac{\sin(\phi_1/\sqrt{a})}{\sin(\phi_1)} \right)^{n-2} \cos^2 \phi_1 \sin^{n-2} \phi_1 d\phi_1 \\ &= \int_0^\pi V'(\phi_1/\sqrt{a})^2 a^{(2n-3)/2} \sin^{n-2}(\phi_1/\sqrt{a}) \cos^2 \phi_1 d\phi_1 \\ &= \int_0^{\pi/\sqrt{a}} V'(\omega_1)^2 a^{n-1} \sin^{n-2} \omega_1 \cos^2(\omega_1 \sqrt{a}) d\omega_1 \\ &\leq a^{n-1} \int_0^{\pi/2} V'(\omega_1)^2 \sin^{n-2} \omega_1 d\omega_1. \end{aligned}$$

The derivative of  $V$  is given by

$$V'(t) = \begin{cases} -2C_v \cos(t) \sin(t), & 0 \leq t \leq \pi/2, \\ 0, & \text{otherwise,} \end{cases} \quad (47)$$

where  $C_v = \sqrt{\frac{2(n+2)n}{3S_0}}$ . Using these formulas and ([1], eqn. 4.3.127), a straightforward calculation shows that

$$\int_0^{\pi/2} V'(\omega_1)^2 \sin^{n-2} \omega_1 d\omega_1 = \frac{4n}{3S_0} \left( 1 - \frac{1}{n} \right) \int_0^\pi \sin^{n-2} \omega_1 d\omega_1. \quad (48)$$

So we have

$$\|U_1\|^2 \leq a^{n-1} \cdot \frac{4n}{3S_0} \left( 1 - \frac{1}{n} \right) \int_0^\pi \sin^{n-2} \omega_1 d\omega_1.$$

Combining with  $I_2$ , we get

$$0 \leq \|U_1\|^2 I_2 \leq a^{n-1} \cdot \frac{4n}{3S_0} \left( 1 - \frac{1}{n} \right) \cdot \frac{1}{n-1} \int_{S^{n-1}} d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) = \frac{4}{3} a^{n-1}. \quad (49)$$

Next we evaluate  $\|U_2\|^2$ . The derivative of  $\Lambda_a$  is given by:

$$\Lambda'_a(\phi_1) = a^{(2n-1)/4} \binom{n-2}{2} \left( \frac{\sin(\phi_1/\sqrt{a})}{\sin(\phi_1)} \right)^{(n-4)/2} \left( \frac{\cos(\phi_1/\sqrt{a})}{\sqrt{a} \sin(\phi_1)} - \frac{\sin(\phi_1/\sqrt{a}) \cos(\phi_1)}{\sin^2(\phi_1)} \right), \quad (50)$$

hence

$$U_2(\phi_1) = V(\phi_1/\sqrt{a})a^{(2n-1)/4}(\frac{n-2}{2})\sin^{(n-4)/2}(\phi_1/\sqrt{a})\left(\frac{\cos(\phi_1/\sqrt{a})}{\sqrt{a}} - \frac{\sin(\phi_1/\sqrt{a})\cos(\phi_1)}{\sin(\phi_1)}\right)\cos\phi_1,$$

and

$$\begin{aligned}\|U_2\|^2 &= \int_0^\pi V(\phi_1/\sqrt{a})^2 a^{(2n-1)/2}(\frac{n-2}{2})^2 \sin^{n-4}(\phi_1/\sqrt{a})\left(\frac{\cos(\phi_1/\sqrt{a})}{\sqrt{a}} - \frac{\sin(\phi_1/\sqrt{a})\cos(\phi_1)}{\sin(\phi_1)}\right)^2 \cos^2\phi_1 d\phi_1 \\ &= \int_0^{\pi/\sqrt{a}} V(\omega_1)^2 a^n(\frac{n-2}{2})^2 \sin^{n-4}(\omega_1)\left(\frac{\cos(\omega_1)}{\sqrt{a}} - \frac{\sin(\omega_1)\cos(\omega_1\sqrt{a})}{\sin(\omega_1\sqrt{a})}\right)^2 \cos^2(\omega_1\sqrt{a})d\omega_1.\end{aligned}$$

Recall that  $V$  is supported on  $[0, \pi/2]$ . For  $\omega_1$  in this range, we have the following crude bound: (using [1], eqn. 4.3.81)

$$\left|\frac{\cos(\omega_1)}{\sqrt{a}} - \frac{\sin(\omega_1)\cos(\omega_1\sqrt{a})}{\sin(\omega_1\sqrt{a})}\right| \leq \frac{1}{\sqrt{a}} + \frac{\sin(\omega_1)}{\sin(\omega_1\sqrt{a})} \cdot \frac{\sin(\omega_1\sqrt{a})}{\omega_1\sqrt{a}} \leq \frac{1}{\sqrt{a}} + \frac{1}{\sqrt{a}}.$$

Also, we have  $\cos^2(\omega_1\sqrt{a}) \leq 1$ . Hence,

$$\begin{aligned}\|U_2\|^2 &\leq \int_0^{\pi/2} V(\omega_1)^2 a^n(\frac{n-2}{2})^2 \sin^{n-4}(\omega_1)(\frac{2}{\sqrt{a}})^2 d\omega_1 \\ &= 4a^{n-1}(\frac{n-2}{2})^2 \int_0^{\pi/2} V(\omega_1)^2 \sin^{n-4}\omega_1 d\omega_1.\end{aligned}$$

Note that

$$\int_0^{\pi/2} V(\omega_1)^2 \sin^{n-4}\omega_1 d\omega_1 = \left(1 + \frac{5}{n-3}\right) \int_0^{\pi/2} V(\omega_1)^2 \sin^{n-2}\omega_1 d\omega_1, \quad (51)$$

using the definition of  $V$ , and integration by parts.

So we have:

$$\|U_2\|^2 \leq 4a^{n-1}(\frac{n-2}{2})^2(1 + \frac{5}{n-3}) \int_0^{\pi/2} V(\omega_1)^2 \sin^{n-2}\omega_1 d\omega_1.$$

Combining with  $I_2$ , we get:

$$\begin{aligned}0 \leq \|U_2\|^2 I_2 &\leq 4a^{n-1}(\frac{n-2}{2})^2(1 + \frac{5}{n-3})\frac{1}{n-1} \int_{S^{n-1}} V(\omega_1)^2 d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) \\ &\leq a^{n-1}(n-2)(1 + \frac{5}{n-3}).\end{aligned} \quad (52)$$

So, by substituting into (46), we have

$$0 \leq I_{C1}I_2 \leq 2 \cdot \frac{4}{3}a^{n-1} + 2 \cdot a^{n-1}(n-2)(1 + \frac{5}{n-3}). \quad (53)$$

Finally, we turn to  $I_{Cr}$ . Combining it with  $I_{C1}$  and  $I_2$ , we have

$$0 \leq \int_0^\eta I_{Cr}I_{C1}I_2 \frac{da}{a^{n+1}} \leq \left(\frac{8}{3} + 2(n-2)(1 + \frac{5}{n-3})\right) \int_0^\eta I_{Cr} \frac{da}{a^2}.$$

We can bound the integral on the right hand side as follows.

$$\begin{aligned}\int_0^\eta I_{Cr} \frac{da}{a^2} &= \int_0^\eta \int_0^\infty F_0(r)^2 W(\lambda ar)^2 r^{n-3} dr \frac{da}{a^2} \\ &= \int_0^\infty F_0(r)^2 \int_0^\eta W(\lambda ar)^2 \frac{da}{a^2} r^{n-3} dr.\end{aligned}$$

Using the fact that  $W$  is supported on  $[1/e, 1]$ ,

$$\int_0^\eta W(\lambda ar)^2 \frac{da}{a^2} = \int_0^{\lambda \eta r} W(\alpha)^2 \frac{d\alpha}{\alpha^2} \lambda r \leq \int_0^{\lambda \eta r} W(\alpha)^2 \frac{d\alpha}{\alpha} e \lambda r \leq e \lambda r, \quad (54)$$

and vanishes when  $r \leq 1/(\lambda \eta e)$ . Hence

$$\int_0^\eta I_{Cr} \frac{da}{a^2} \leq e \lambda \int_{1/(\lambda \eta e)}^\infty F_0(r)^2 r^{n-2} dr.$$

$$0 \leq \int_0^\eta I_{Cr} I_{C1} I_2 \frac{da}{a^{n+1}} \leq \left( \frac{8}{3} + 2(n-2)(1 + \frac{5}{n-3}) \right) e \lambda \int_{1/(\lambda \eta e)}^\infty F_0(r)^2 r^{n-2} dr. \quad (55)$$

## B.2 The variance of $\vec{b}$ parallel to $\vec{\theta}$

Finally, we seek to bound the variance of  $\vec{b}$ , in the direction parallel to  $\vec{\theta}$ . The analysis is similar to the previous case (i.e., the variance of  $\vec{b}$  orthogonal to  $\vec{\theta}$ ).

The variance of  $\vec{b}$  parallel to  $\vec{\theta}$  is:

$$\mathbb{E}((\vec{b} \cdot \vec{\theta})^2) = \int_{S^{n-1}} \int_0^1 \int_{\mathbb{R}^n} (\vec{b} \cdot \vec{\theta})^2 |\Gamma_f(a, \vec{b}, \vec{\theta})|^2 d\vec{b} \frac{da}{a^{n+1}} d\sigma(\vec{\theta}). \quad (56)$$

We can simplify this by taking advantage of rotational symmetry. Fix a vector  $\vec{u} = (1, 0, \dots, 0)$ , and for each  $\vec{\theta}$ , let  $R$  be a rotation that maps  $\vec{\theta}$  to  $\vec{u}$ . Then

$$\mathbb{E}((\vec{b} \cdot \vec{\theta})^2) = \int_{S^{n-1}} \int_0^1 \int_{\mathbb{R}^n} (R(\vec{b}) \cdot \vec{u})^2 |\Gamma_f(a, R(\vec{b}), \vec{u})|^2 d\vec{b} \frac{da}{a^{n+1}} d\sigma(\vec{\theta}). \quad (57)$$

Then change variables  $\vec{b} \mapsto R^{-1}(\vec{b})$ . The integrand is now independent of  $\vec{\theta}$ , so we can do the  $\vec{\theta}$  integral. We get:

$$\mathbb{E}((\vec{b} \cdot \vec{\theta})^2) = S_0 \int_0^1 \int_{\mathbb{R}^n} b_1^2 |\Gamma_f(a, \vec{b}, \vec{u})|^2 d\vec{b} \frac{da}{a^{n+1}}. \quad (58)$$

We now introduce some new notation,

$$\Phi_{a, \vec{\theta}}(\vec{b}) := \Gamma_f(a, \vec{b}, \vec{\theta}), \quad (59)$$

to emphasize that we view this as a function of  $\vec{b}$ . By equation (1), the Fourier transform of  $\Phi_{a, \vec{\theta}}$  is given by

$$\hat{\Phi}_{a, \vec{\theta}}(\vec{k}) = \hat{f}(\vec{k}) \chi_{a, \vec{\theta}}(\vec{k}). \quad (60)$$

And we have, by Plancherel's theorem:

$$\begin{aligned} \mathbb{E}((\vec{b} \cdot \vec{\theta})^2) &= S_0 \int_0^1 \int_{\mathbb{R}^n} b_1^2 |\Phi_{a, \vec{u}}(\vec{b})|^2 d\vec{b} \frac{da}{a^{n+1}} \\ &= S_0 \int_0^1 \int_{\mathbb{R}^n} \left| \frac{1}{2\pi i} \frac{\partial}{\partial k_1} \hat{\Phi}_{a, \vec{u}}(\vec{k}) \right|^2 d\vec{k} \frac{da}{a^{n+1}}. \end{aligned} \quad (61)$$

Now, using spherical coordinates  $\vec{k} = (r, \phi_1, \dots, \phi_{n-1})$ , we write  $\hat{\Phi}_{a, \vec{u}}(\vec{k})$  as a product of a radial part and an angular part:

$$\hat{\Phi}_{a, \vec{u}}(\vec{k}) = L(r) M(\phi_1), \quad (62)$$

where

$$L(r) = F_0(r)W(\lambda ar), \quad M(\phi_1) = V(\phi_1/\sqrt{a})\Lambda_a(\phi_1). \quad (63)$$

Then we have

$$\frac{\partial}{\partial k_1} \hat{\Phi}_{a,\vec{u}}(\vec{k}) = L'(r)M(\phi_1)\frac{\partial r}{\partial k_1} + L(r)M'(\phi_1)\frac{\partial \phi_1}{\partial k_1}, \quad (64)$$

where

$$\frac{\partial r}{\partial k_1} = \cos \phi_1, \quad \frac{\partial \phi_1}{\partial k_1} = -\frac{\sin \phi_1}{r}. \quad (65)$$

Now we can expand out the following integral: (note that  $\hat{\Phi}_{a,\vec{u}}(\vec{k})$  is real)

$$\begin{aligned} \int_{\mathbb{R}^n} \left| \frac{\partial}{\partial k_1} \hat{\Phi}_{a,\vec{u}}(\vec{k}) \right|^2 d\vec{k} &= \int_{S^{n-1}} \int_0^\infty \left( L'(r)M(\phi_1) \cos \phi_1 \right. \\ &\quad \left. - L(r)M'(\phi_1)r^{-1} \sin \phi_1 \right)^2 r^{n-1} dr d\sigma(\vec{\phi}) \\ &= K_{Ar}K_{A1}K_2 - 2K_{Br}K_{B1}K_2 + K_{Cr}K_{C1}K_2, \end{aligned} \quad (66)$$

where we define

$$K_{Ar} = \int_0^\infty L'(r)^2 r^{n-1} dr \quad (67)$$

$$K_{A1} = \int_0^\pi M(\phi_1)^2 \cos^2 \phi_1 \sin^{n-2} \phi_1 d\phi_1 \quad (68)$$

$$K_2 = \int_{S^{n-2}} d\sigma(\phi_2, \dots, \phi_{n-1}) \quad (69)$$

$$K_{Br} = \int_0^\infty L'(r)L(r)r^{n-2} dr \quad (70)$$

$$K_{B1} = \int_0^\pi M'(\phi_1)M(\phi_1) \cos \phi_1 \sin^{n-1} \phi_1 d\phi_1 \quad (71)$$

$$K_{Cr} = \int_0^\infty L(r)^2 r^{n-3} dr \quad (72)$$

$$K_{C1} = \int_0^\pi M'(\phi_1)^2 \sin^n \phi_1 d\phi_1. \quad (73)$$

Thus we can write the variance of  $\vec{b}$  parallel to  $\vec{\theta}$  as:

$$\mathbb{E}((\vec{b} \cdot \vec{\theta})^2) = \frac{S_0}{(2\pi)^2} \int_0^1 \left( K_{Ar}K_{A1}K_2 - 2K_{Br}K_{B1}K_2 + K_{Cr}K_{C1}K_2 \right) \frac{da}{a^{n+1}}. \quad (74)$$

A similar formula gives the variance conditioned on observing  $a \leq \eta$ :

$$\begin{aligned} \mathbb{E}((\vec{b} \cdot \vec{\theta})^2 \mid a \leq \eta) \\ = \frac{1}{\Pr[a \leq \eta]} \frac{S_0}{(2\pi)^2} \int_0^\eta \left( K_{Ar}K_{A1}K_2 - 2K_{Br}K_{B1}K_2 + K_{Cr}K_{C1}K_2 \right) \frac{da}{a^{n+1}}. \end{aligned} \quad (75)$$

We would then like to bound the various integrals appearing on the right hand side.

### B.2.1

We begin with the integral  $\int_0^\eta K_{Br} K_{B1} K_2 da/a^{n+1}$ .

Note that  $K_2 = (n-1)I_2$ , while  $K_{B1} = I_{B1}$  and  $K_{Br} = I_{Br}$ . Using the argument from the previous section, we have:

$$\left| \int_0^\eta K_{Br} K_{B1} K_2 \frac{da}{a^{n+1}} \right| \leq \frac{1}{4}(n-1)(n-2) \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-3} dr. \quad (76)$$

### B.2.2

Next, consider the integral  $\int_0^\eta K_{Ar} K_{A1} K_2 da/a^{n+1}$ . We already have a bound for  $K_2$ . For  $K_{A1}$ , we write:

$$\begin{aligned} K_{A1} &= \int_0^\pi V(\phi_1/\sqrt{a})^2 \Lambda_a(\phi_1)^2 \cos^2 \phi_1 \sin^{n-2} \phi_1 d\phi_1 \\ &= \int_0^\pi V(\phi_1/\sqrt{a})^2 a^{(2n-1)/2} \left( \frac{\sin(\phi_1/\sqrt{a})}{\sin \phi_1} \right)^{n-2} \cos^2 \phi_1 \sin^{n-2} \phi_1 d\phi_1 \\ &= \int_0^{\pi/\sqrt{a}} V(\omega_1)^2 a^n \sin^{n-2} \omega_1 \cos^2(\omega_1 \sqrt{a}) d\omega_1. \end{aligned}$$

Using the fact that  $V$  is supported on  $[0, \pi/2]$ , and the simple bound  $\cos^2(\omega_1 \sqrt{a}) \leq 1$ , we get:

$$0 \leq K_{A1} \leq \int_0^{\pi/2} V(\omega_1)^2 a^n \sin^{n-2} \omega_1 d\omega_1.$$

Combining with  $K_2$ , we get:

$$0 \leq K_{A1} K_2 \leq a^n \int_{S^{n-1}} V(\omega_1)^2 d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) = a^n. \quad (77)$$

We now turn to  $K_{Ar}$ . First, combining with  $K_{A1}$  and  $K_2$ , we have

$$0 \leq \int_0^\eta K_{Ar} K_{A1} K_2 \frac{da}{a^{n+1}} \leq \int_0^\eta K_{Ar} \frac{da}{a}. \quad (78)$$

Note that  $K_{Ar} = I_{Ar}$ , so we can upper-bound  $K_{Ar}$  as in the previous section:

$$0 \leq K_{Ar} \leq 2\|G_1\|^2 + 2\|G_2\|^2, \quad (79)$$

where we define

$$G_1(r) = F'_0(r) W(\lambda ar) r^{(n-1)/2} \quad (80)$$

$$G_2(r) = F_0(r) W'(\lambda ar) \lambda ar^{(n-1)/2}. \quad (81)$$

Thus we have

$$0 \leq \int_0^\eta K_{Ar} K_{A1} K_2 \frac{da}{a^{n+1}} \leq 2 \left( \int_0^\eta \|G_1\|^2 \frac{da}{a} + \int_0^\eta \|G_2\|^2 \frac{da}{a} \right). \quad (82)$$

We then want to upper-bound the integrals  $\int_0^\eta \|G_1\|^2 da/a$  and  $\int_0^\eta \|G_2\|^2 da/a$ .

For the first one, we have:

$$\begin{aligned}\int_0^\eta \|G_1\|^2 \frac{da}{a} &= \int_0^\eta \int_0^\infty F_0'(r)^2 W(\lambda ar)^2 r^{n-1} dr \frac{da}{a} \\ &= \int_0^\infty F_0'(r)^2 \int_0^\eta W(\lambda ar)^2 \frac{da}{a} r^{n-1} dr.\end{aligned}$$

Using equation (29), we get:

$$\boxed{\int_0^\eta \|G_1\|^2 \frac{da}{a} \leq \int_{1/(\lambda\eta e)}^\infty F_0'(r)^2 r^{n-1} dr.} \quad (83)$$

For the second integral, we have:

$$\begin{aligned}\int_0^\eta \|G_2\|^2 \frac{da}{a} &= \int_0^\eta \int_0^\infty F_0(r)^2 W'(\lambda ar)^2 \lambda^2 a^2 r^{n-1} dr \frac{da}{a} \\ &= \int_0^\infty F_0(r)^2 \int_0^\eta W'(\lambda ar)^2 \lambda^2 a da r^{n-1} dr.\end{aligned}$$

Recall that the derivative of  $W$  is given by equation (40). So we can write

$$\int_0^\eta W'(\lambda ar)^2 \lambda^2 a da = \int_0^{\lambda\eta r} W'(\alpha)^2 \alpha d\alpha \frac{1}{r^2} \leq \int_{1/e}^1 C_w^2 (\pi/\alpha)^2 \alpha d\alpha \frac{1}{r^2} = \frac{8}{3} \pi^2 \frac{1}{r^2}, \quad (84)$$

and vanishes when  $r \leq 1/(\lambda\eta e)$ . Hence,

$$\boxed{\int_0^\eta \|G_2\|^2 \frac{da}{a} \leq \frac{8}{3} \pi^2 \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-3} dr.} \quad (85)$$

### B.2.3

Finally, we consider the integral  $\int_0^\eta K_{Cr} K_{C1} K_2 da / a^{n+1}$ . We already have a bound for  $K_2$ . For  $K_{C1}$  we can write:

$$\begin{aligned}0 \leq K_{C1} &= \int_0^\pi \left( V'(\phi_1/\sqrt{a})(1/\sqrt{a})\Lambda_a(\phi_1) + V(\phi_1/\sqrt{a})\Lambda'_a(\phi_1) \right)^2 \sin^n \phi_1 d\phi_1 \\ &= \|\tilde{U}_1 + \tilde{U}_2\|^2 \leq 2\|\tilde{U}_1\|^2 + 2\|\tilde{U}_2\|^2,\end{aligned} \quad (86)$$

where we define

$$\tilde{U}_1(\phi_1) = V'(\phi_1/\sqrt{a})(1/\sqrt{a})\Lambda_a(\phi_1) \sin^{n/2} \phi_1 \quad (87)$$

$$\tilde{U}_2(\phi_1) = V(\phi_1/\sqrt{a})\Lambda'_a(\phi_1) \sin^{n/2} \phi_1. \quad (88)$$

Then

$$0 \leq K_{C1} K_2 \leq 2\|\tilde{U}_1\|^2 K_2 + 2\|\tilde{U}_2\|^2 K_2. \quad (89)$$

We now evaluate  $\|\tilde{U}_1\|^2$  and  $\|\tilde{U}_2\|^2$ .

For  $\|\tilde{U}_1\|^2$ , we can write

$$\begin{aligned}
\|\tilde{U}_1\|^2 &= \int_0^\pi V'(\phi_1/\sqrt{a})^2 a^{-1} a^{(2n-1)/2} \left( \frac{\sin(\phi_1/\sqrt{a})}{\sin(\phi_1)} \right)^{n-2} \sin^n \phi_1 d\phi_1 \\
&= \int_0^\pi V'(\phi_1/\sqrt{a})^2 a^{(2n-3)/2} \sin^{n-2}(\phi_1/\sqrt{a}) \sin^2 \phi_1 d\phi_1 \\
&= \int_0^{\pi/\sqrt{a}} V'(\omega_1)^2 a^{n-1} \sin^{n-2} \omega_1 \sin^2(\omega_1 \sqrt{a}) d\omega_1 \\
&\leq \frac{\pi^2}{4} a^n \int_0^{\pi/2} V'(\omega_1)^2 \sin^{n-2} \omega_1 d\omega_1.
\end{aligned}$$

(In the last step we used the bound  $\sin^2(\omega_1 \sqrt{a}) \leq \omega_1^2 a \leq \frac{\pi^2}{4} a$ .)

Then, using equation (48), we have

$$\|\tilde{U}_1\|^2 \leq \frac{\pi^2}{4} a^n \cdot \frac{4n}{3S_0} \left(1 - \frac{1}{n}\right) \int_0^\pi \sin^{n-2} \omega_1 d\omega_1.$$

Combining with  $K_2$ , we get

$$0 \leq \|\tilde{U}_1\|^2 K_2 \leq \frac{\pi^2}{4} a^n \cdot \frac{4n}{3S_0} \left(1 - \frac{1}{n}\right) \cdot \int_{S^{n-1}} d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) = \frac{\pi^2}{3} a^n (n-1). \quad (90)$$

Next we evaluate  $\|\tilde{U}_2\|^2$ . The derivative of  $\Lambda_a$  is given by equation (50), hence

$$\tilde{U}_2(\phi_1) = V(\phi_1/\sqrt{a}) a^{(2n-1)/4} \left(\frac{n-2}{2}\right) \sin^{(n-4)/2}(\phi_1/\sqrt{a}) \left( \frac{\cos(\phi_1/\sqrt{a})}{\sqrt{a}} - \frac{\sin(\phi_1/\sqrt{a}) \cos(\phi_1)}{\sin(\phi_1)} \right) \sin \phi_1,$$

and

$$\begin{aligned}
\|\tilde{U}_2\|^2 &= \int_0^\pi V(\phi_1/\sqrt{a})^2 a^{(2n-1)/2} \left(\frac{n-2}{2}\right)^2 \sin^{n-4}(\phi_1/\sqrt{a}) \left( \frac{\cos(\phi_1/\sqrt{a})}{\sqrt{a}} - \frac{\sin(\phi_1/\sqrt{a}) \cos(\phi_1)}{\sin(\phi_1)} \right)^2 \sin^2 \phi_1 d\phi_1 \\
&= \int_0^{\pi/\sqrt{a}} V(\omega_1)^2 a^n \left(\frac{n-2}{2}\right)^2 \sin^{n-4}(\omega_1) \left( \frac{\cos(\omega_1)}{\sqrt{a}} - \frac{\sin(\omega_1) \cos(\omega_1 \sqrt{a})}{\sin(\omega_1 \sqrt{a})} \right)^2 \sin^2(\omega_1 \sqrt{a}) d\omega_1.
\end{aligned}$$

Recall that  $V$  is supported on  $[0, \pi/2]$ . For  $\omega_1$  in this range, we have the following crude bound: (using [1], eqn. 4.3.81)

$$\left| \frac{\cos(\omega_1)}{\sqrt{a}} - \frac{\sin(\omega_1) \cos(\omega_1 \sqrt{a})}{\sin(\omega_1 \sqrt{a})} \right| \leq \frac{1}{\sqrt{a}} + \frac{\sin(\omega_1)}{\sin(\omega_1 \sqrt{a})} \cdot \frac{\sin(\omega_1 \sqrt{a})}{\omega_1 \sqrt{a}} \leq \frac{1}{\sqrt{a}} + \frac{1}{\sqrt{a}}.$$

Also, we have  $\sin^2(\omega_1 \sqrt{a}) \leq \omega_1^2 a \leq \frac{\pi^2}{4} a$ . Hence,

$$\begin{aligned}
\|\tilde{U}_2\|^2 &\leq \int_0^{\pi/2} V(\omega_1)^2 a^n \left(\frac{n-2}{2}\right)^2 \sin^{n-4}(\omega_1) \left(\frac{2}{\sqrt{a}}\right)^2 \frac{\pi^2}{4} a d\omega_1 \\
&= \pi^2 a^n \left(\frac{n-2}{2}\right)^2 \int_0^{\pi/2} V(\omega_1)^2 \sin^{n-4} \omega_1 d\omega_1.
\end{aligned}$$

By equation (51), we have:

$$\|\tilde{U}_2\|^2 \leq \pi^2 a^n \left(\frac{n-2}{2}\right)^2 \left(1 + \frac{5}{n-3}\right) \int_0^{\pi/2} V(\omega_1)^2 \sin^{n-2} \omega_1 d\omega_1.$$

Combining with  $K_2$ , we get:

$$\begin{aligned} 0 \leq \|\tilde{U}_2\|^2 K_2 &\leq \pi^2 a^n \left(\frac{n-2}{2}\right)^2 \left(1 + \frac{5}{n-3}\right) \int_{S^{n-1}} V(\omega_1)^2 d\sigma(\omega_1, \phi_2, \dots, \phi_{n-1}) \\ &\leq \frac{\pi^2}{4} a^n (n-2)^2 \left(1 + \frac{5}{n-3}\right). \end{aligned} \quad (91)$$

So, by substituting into (89), we have

$$0 \leq K_{C1} K_2 \leq 2 \cdot \frac{\pi^2}{3} a^n (n-1) + 2 \cdot \frac{\pi^2}{4} a^n (n-2)^2 \left(1 + \frac{5}{n-3}\right). \quad (92)$$

Finally, we turn to  $K_{Cr}$ . Combining it with  $K_{C1}$  and  $K_2$ , we have

$$0 \leq \int_0^\eta K_{Cr} K_{C1} K_2 \frac{da}{a^{n+1}} \leq \left( \frac{2\pi^2}{3} (n-1) + \frac{\pi^2}{2} (n-2)^2 \left(1 + \frac{5}{n-3}\right) \right) \int_0^\eta K_{Cr} \frac{da}{a}.$$

We can bound the integral on the right hand side as follows.

$$\begin{aligned} \int_0^\eta K_{Cr} \frac{da}{a} &= \int_0^\eta \int_0^\infty F_0(r)^2 W(\lambda ar)^2 r^{n-3} dr \frac{da}{a} \\ &= \int_0^\infty F_0(r)^2 \int_0^\eta W(\lambda ar)^2 \frac{da}{a} r^{n-3} dr. \end{aligned}$$

Using equation (29), we get

$$\int_0^\eta K_{Cr} \frac{da}{a} \leq \int_{1/(\lambda \eta e)}^\infty F_0(r)^2 r^{n-3} dr.$$

$$0 \leq \int_0^\eta K_{Cr} K_{C1} K_2 \frac{da}{a^{n+1}} \leq \left( \frac{2\pi^2}{3} (n-1) + \frac{\pi^2}{2} (n-2)^2 \left(1 + \frac{5}{n-3}\right) \right) \int_{1/(\lambda \eta e)}^\infty F_0(r)^2 r^{n-3} dr. \quad (93)$$

## C The Ball in $\mathbb{R}^n$

We prove Theorem 3.

### C.1 The low-frequency components

First, we claim that almost all the power in  $\hat{f}$  is located at frequencies above some threshold  $1/\lambda$ . This justifies our use of the curvelet transform, and theorems 1 and 2, for an appropriate choice of the parameter  $\lambda$ .

We start by proving an upper-bound on the integral  $\int_0^z t^{-1} J_\nu(t)^2 dt$ , for  $\nu > 0$ . Note that ([1], eqn. (9.1.62))

$$|J_\nu(t)| \leq \frac{(\frac{1}{2}t)^\nu}{\nu!} \quad (\nu \geq -\frac{1}{2}, t \geq 0). \quad (94)$$

Also ([1], eqn. (6.1.38)),

$$\nu! > \sqrt{2\pi\nu} \nu^{(\nu+\frac{1}{2})} e^{-\nu} \quad (\nu > 0). \quad (95)$$

Hence

$$|J_\nu(t)| < \frac{(\frac{1}{2}t)^\nu}{\sqrt{2\pi\nu} \nu^{(\nu+\frac{1}{2})} e^{-\nu}} = \frac{1}{\sqrt{2\pi\nu}} \left(\frac{te}{2\nu}\right)^\nu, \quad (96)$$



and

$$\int_0^z t^{-1} J_\nu(t)^2 dt < \int_0^z \frac{1}{t} \frac{1}{2\pi\nu} \left(\frac{te}{2\nu}\right)^{2\nu} dt = \frac{1}{2\pi\nu} \left(\frac{e}{2\nu}\right)^{2\nu} \frac{1}{2\nu} t^{2\nu} \Big|_0^z = \frac{1}{4\pi\nu^2} \left(\frac{ez}{2\nu}\right)^{2\nu}. \quad (97)$$

This upper bound is useful when  $z \leq 2\nu/e$ .

We can now calculate the amount of power contained in the low-frequency components of  $f$ :

$$\begin{aligned} \int_{|\vec{k}| \leq z} |\hat{f}(\vec{k})|^2 d\vec{k} &= \int_0^z \frac{n}{S_0} \frac{1}{\rho^n} J_{n/2}(2\pi\rho\beta)^2 \cdot S_0 \rho^{n-1} d\rho \\ &= n \int_0^{2\pi\beta z} t^{-1} J_{n/2}(t)^2 \cdot dt \\ &< n \frac{1}{\pi n^2} \left(\frac{2\pi\beta ez}{n}\right)^n. \end{aligned} \quad (98)$$

Setting  $z = n/(2\pi\beta e)$ , we get

$$\int_{|\vec{k}| \leq n/(2\pi\beta e)} |\hat{f}(\vec{k})|^2 d\vec{k} < \frac{1}{\pi n}. \quad (99)$$

Recall that we set the parameter  $\lambda$  so that  $\lambda \geq 2\pi\beta e/n$ . So the region  $\{|\vec{k}| \leq 1/\lambda\}$  contains at most a  $1/(\pi n)$  fraction of the total power.

## C.2 The decay of $J_\nu(x)$

We now prove some technical lemmas on the decay of  $J_\nu(x)$  for  $x \geq 2\nu$ ,  $\nu \geq 1/2$ . These follow from classical results on Bessel functions [1, 30], though some care is required near the transition region at  $x \approx \nu$ . In particular, the usual asymptotic expansions for  $J_\nu(x)$  only work when  $x \geq \nu^2$ , or when  $x = \alpha\nu$  for some fixed constant  $\alpha$ . For our purposes, we use an asymptotic expansion of  $J_\nu(x)^2 + Y_\nu(x)^2$ , that behaves well when  $x \geq \nu$ .

### C.2.1

We start by quoting the following result from ([30], p.447). Define

$$M_\nu(x) = \sqrt{J_\nu(x)^2 + Y_\nu(x)^2}. \quad (100)$$

Then for all  $x \geq \nu \geq 1/2$ ,

$$\frac{2}{\pi x} < M_\nu(x)^2 < \frac{2}{\pi \sqrt{x^2 - \nu^2}}. \quad (101)$$

This immediately implies an upper bound on  $J_\nu(x)^2$ , for all  $x \geq 2\nu$ ,  $\nu \geq 1/2$ :

$$J_\nu(x)^2 \leq M_\nu(x)^2 < \frac{2}{\pi \sqrt{x^2 - \nu^2}} \leq \frac{2}{\pi x} \cdot \frac{2}{\sqrt{3}}. \quad (102)$$

### C.2.2

We next prove a lower bound on  $|J_\nu(x)|$ , for  $x$  within certain intervals. Note that  $J_\nu(x)$  is large at a zero of  $Y_\nu(x)$ . We will show that (1) the zeroes of  $Y_\nu(x)$  are not too far apart, and (2)  $J_\nu(x)$  is large in a neighborhood around each zero of  $Y_\nu(x)$ .

To see this, note that  $J_\nu(x)$  and  $Y_\nu(x)$  can be written in terms of a modulus and phase,

$$J_\nu(x) = M_\nu(x) \cos \theta_\nu(x), \quad (103)$$

$$Y_\nu(x) = M_\nu(x) \sin \theta_\nu(x), \quad (104)$$

where  $M_\nu(x)$  is as defined above, and  $\theta_\nu(x)$  satisfies the equation

$$\theta'_\nu(x) = \frac{2}{\pi x M_\nu(x)^2} \quad (105)$$

(see [1], eqn. 9.2.21, and [30], p.514). This implies lower and upper bounds on  $\theta'_\nu(x)$ , for all  $x \geq 2\nu$ :

$$\frac{\sqrt{3}}{2} < \theta'_\nu(x) < 1. \quad (106)$$

First, we claim that for any  $t \geq 2\nu$ , the interval  $[t, t + \frac{2\pi}{\sqrt{3}}]$  contains a zero of  $Y_\nu(x)$ . To see this, write the following, for any  $\delta \geq 0$ :

$$\theta_\nu(x + \delta) = \theta_\nu(x) + \int_x^{x+\delta} \theta'_\nu(y) dy \geq \theta_\nu(x) + \frac{\sqrt{3}}{2} \delta.$$

So  $\theta_\nu(t + \frac{2\pi}{\sqrt{3}}) \geq \theta_\nu(t) + \pi$ . So  $\theta_\nu(x)$  must equal an integer multiple of  $\pi$  for some  $x \in [t, t + \frac{2\pi}{\sqrt{3}}]$ ; and  $Y_\nu(x)$  must vanish at that point. This proves our first claim.

Second, let  $\phi$  be a zero of  $Y_\nu(x)$ , satisfying  $\phi \geq 2\nu$ . We claim that, for any  $\delta \in [-\pi/2, \pi/2]$ ,

$$|J_\nu(\phi + \delta)| \geq M_\nu(\phi + \delta) \cos \delta. \quad (107)$$

To see this, write

$$|J_\nu(\phi + \delta)| = M_\nu(\phi + \delta) |\cos \theta_\nu(\phi + \delta)| = M_\nu(\phi + \delta) |\cos |\theta_\nu(\phi + \delta) - \theta_\nu(\phi)||.$$

(The last step follows because  $\theta_\nu(\phi)$  is an integer multiple of  $\pi$ .) Then note that

$$|\theta_\nu(\phi + \delta) - \theta_\nu(\phi)| = \left| \int_\phi^{\phi+\delta} \theta'_\nu(y) dy \right| \leq |\delta|.$$

Hence we have

$$|\cos |\theta_\nu(\phi + \delta) - \theta_\nu(\phi)|| \geq |\cos |\delta|| = \cos \delta.$$

This proves our second claim.

### C.2.3

Finally, we prove the following lower bound on a sum of squares of Bessel functions:

Let  $\nu_1, \dots, \nu_m \in [1/2, \nu_{\max}]$ . Let  $t \geq 2\nu_{\max}$ . Then there exists some  $t' \in [t - \frac{\pi}{2}, t + \frac{2\pi}{\sqrt{3}} + \frac{\pi}{2}]$  such that

$$\sum_{k=1}^m J_{\nu_k}(t')^2 \geq \frac{m}{7t'}. \quad (108)$$

Proof: Essentially, we will show that there must exist a point  $t'$  where a constant fraction of the functions  $J_{\nu_k}(t')^2$  ( $k = 1, \dots, m$ ) are large simultaneously.

Define the interval  $I = [t, t + \frac{2\pi}{\sqrt{3}}]$ . For each  $k = 1, \dots, m$ ,  $I$  contains a zero of  $Y_{\nu_k}(x)$ , call it  $\phi_k$ . Now define the function

$$\chi_k(x) = \begin{cases} \cos^2(x - \phi_k) & \text{if } \phi_k - \frac{\pi}{2} \leq x \leq \phi_k + \frac{\pi}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that

$$J_{\nu_k}(x)^2 \geq M_{\nu_k}(x)^2 \chi_k(x) > \frac{2}{\pi x} \chi_k(x).$$

Furthermore, define the function

$$u(x) = \sum_{k=1}^m \chi_k(x),$$

and note that

$$\sum_{k=1}^m J_{\nu_k}(x)^2 \geq \frac{2}{\pi x} u(x).$$

Define the interval  $I' = [t - \frac{\pi}{2}, t + \frac{2\pi}{\sqrt{3}} + \frac{\pi}{2}]$ ; this interval contains the support of all of the functions  $\chi_k(x)$  ( $k = 1, \dots, m$ ). Then write

$$\int_{I'} u(x) dx = \sum_{k=1}^m \int_{\phi_k - (\pi/2)}^{\phi_k + (\pi/2)} \chi_k(x) dx = m \cdot \frac{\pi}{2}.$$

So there must exist a point  $t' \in I'$  such that

$$u(t') \geq \frac{1}{|I'|} \int_{I'} u(x) dx \geq \frac{1}{7} \cdot m \cdot \frac{\pi}{2} = \frac{\pi}{14} m,$$

and the claim follows.

### C.3 The probability of observing a scale $a$

Next we claim that  $\hat{f}$  has a heavy tail. This implies that we will observe fine-scale elements ( $a$  small) with significant probability.

Again, we start by proving a lower bound on  $\int_z^\infty t^{-1} J_\nu(t)^2 dt$ , when  $\nu$  is of the form  $m$  or  $m + (1/2)$  (where  $m$  is an integer),  $\nu \geq 1$ , and  $z \geq 2\nu$ . We will show that:

$$\int_z^\infty t^{-1} J_\nu(t)^2 dt \geq \left(1 - \frac{1}{2\nu}\right) \frac{1}{7(z + 5.20)}. \quad (109)$$

First, consider the case of  $\nu = m$ . We assume  $m \geq 1$  and  $z \geq 2m$ . Using ([1], eqn. 11.3.36), we

can write

$$\begin{aligned}
\int_z^\infty t^{-1} J_\nu(t)^2 dt &= \int_z^\infty t^{-1} J_m(t)^2 dt \\
&= -\frac{1}{2m} \left( J_0(t)^2 + J_m(t)^2 + 2 \sum_{k=1}^{m-1} J_k(t)^2 \right) \Big|_z^\infty \\
&= \frac{1}{2m} \left( J_0(z)^2 + J_m(z)^2 + 2 \sum_{k=1}^{m-1} J_k(z)^2 \right) \\
&\geq \frac{1}{2m} \left( J_m(z)^2 + 2 \sum_{k=1}^{m-1} J_k(z)^2 \right).
\end{aligned}$$

Then, using the lemma from the previous section, we get the following, for some  $z' \in [z, z + 5.20]$ :

$$\int_z^\infty t^{-1} J_\nu(t)^2 dt \geq \frac{1}{2m} \frac{2m-1}{7z'} \geq \left(1 - \frac{1}{2\nu}\right) \frac{1}{7(z+5.20)}. \quad (110)$$

Next, consider the case  $\nu = m + (1/2)$ . We assume  $m \geq 1$  and  $z \geq 2m + 1$ . Using ([1], eqn. 11.3.36), we can write

$$\begin{aligned}
\int_z^\infty t^{-1} J_\nu(t)^2 dt &= \int_z^\infty t^{-1} J_{m+(1/2)}(t)^2 dt \\
&= \frac{1}{2m+1} \int_z^\infty t^{-1} J_{1/2}(t)^2 dt - \frac{1}{2m+1} \left( J_{1/2}(t)^2 + J_{m+(1/2)}(t)^2 + 2 \sum_{k=1}^{m-1} J_{k+(1/2)}(t)^2 \right) \Big|_z^\infty \\
&\geq \frac{1}{2m+1} \left( J_{1/2}(z)^2 + J_{m+(1/2)}(z)^2 + 2 \sum_{k=1}^{m-1} J_{k+(1/2)}(z)^2 \right).
\end{aligned}$$

Then, using the lemma from the previous section, we get the following, for some  $z' \in [z, z + 5.20]$ :

$$\int_z^\infty t^{-1} J_\nu(t)^2 dt \geq \frac{1}{2m+1} \frac{2m}{7z'} \geq \left(1 - \frac{1}{2\nu}\right) \frac{1}{7(z+5.20)}. \quad (111)$$

We now proceed to lower-bound the probability of observing a fine-scale element  $a$ . The following bound holds for any  $n \geq 2$  and any  $\eta \leq 1/e^2$ .

$$\begin{aligned}
\Pr[a \leq \eta] &\geq \int_{|\vec{k}| \geq 1/(\lambda\eta)} |\hat{f}(\vec{k})|^2 d\vec{k} \\
&= \int_{1/(\lambda\eta)}^\infty \frac{n}{S_0} \rho^{-n} J_{n/2}(2\pi\rho\beta)^2 \cdot S_0 \rho^{n-1} d\rho \\
&= n \int_{2\pi\beta/(\lambda\eta)}^\infty t^{-1} J_{n/2}(t)^2 dt \\
&\geq n \int_{n/(e\eta)}^\infty t^{-1} J_{n/2}(t)^2 dt,
\end{aligned}$$

where in the last step we used the fact that  $\lambda \geq 2\pi\beta e/n$ . Then, by equation (109), and using the fact that  $n/(e\eta) \geq 2e \geq 5.43$ , we get

$$\Pr[a \leq \eta] \geq n \left(1 - \frac{1}{n}\right) \frac{1}{7(\frac{n}{e\eta} + 5.20)} \geq n \left(1 - \frac{1}{n}\right) \frac{1}{14\frac{n}{e\eta}} = \frac{e\eta}{14} \left(1 - \frac{1}{n}\right). \quad (112)$$

$$\boxed{\Pr[a \leq \eta] \geq \frac{e\eta}{14} \left(1 - \frac{1}{n}\right)}. \quad (113)$$

#### C.4 The variance of $\vec{b}$ orthogonal to $\vec{\theta}$

First, we give a simple upper bound on integrals of the form

$$\int_z^\infty t^{-k} J_\nu(t)^2 dt,$$

for  $k \geq 1$ ,  $z \geq 2\nu$  and  $\nu \geq 1/2$ . This follows from equation (102):

$$\begin{aligned} \int_z^\infty t^{-k} J_\nu(t)^2 dt &\leq \int_z^\infty t^{-k} \frac{2}{\pi t} \frac{2}{\sqrt{3}} dt \\ &= \frac{4}{\pi\sqrt{3}} \int_z^\infty t^{-k-1} dt = \frac{4}{\pi\sqrt{3}} (-1/k) t^{-k} \Big|_z^\infty = \frac{4}{\pi\sqrt{3}k} z^{-k}. \end{aligned} \quad (114)$$

We now use this to upper-bound the integral

$$\int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-k} dr,$$

for  $k \geq 1$  and  $\eta \leq 1/2e^2$ . We write the following: (for the last step, recall that  $\lambda \leq 2 \cdot 2\pi\beta e/n$ , which implies  $2\pi\beta/(\lambda\eta e) \geq n/(2\eta e^2)$ )

$$\begin{aligned} \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-k} dr &= \frac{n}{S_0} \int_{1/(\lambda\eta e)}^\infty J_{n/2}(2\pi\beta r)^2 r^{-k} dr \\ &= \frac{n}{S_0} \int_{2\pi\beta/(\lambda\eta e)}^\infty J_{n/2}(t)^2 t^{-k} dt \cdot (2\pi\beta)^{k-1} \\ &\leq \frac{n}{S_0} \frac{4}{\pi\sqrt{3}k} \left(\frac{2\eta e^2}{n}\right)^k \cdot (2\pi\beta)^{k-1}. \end{aligned} \quad (115)$$

In a similar way, we can upper-bound the integral

$$\int_{1/(\lambda\eta e)}^\infty F'_0(r)^2 r^{n-k} dr,$$

for  $k \geq 1$  and  $\eta \leq \frac{n}{n+2}(1/2e^2)$ . First, note that

$$F'_0(r) = -\sqrt{\frac{n}{S_0}} r^{-n/2} J_{(n/2)+1}(2\pi\beta r)(2\pi\beta). \quad (116)$$

(To see this, write  $F_0(r) = \sqrt{\frac{n}{S_0}} (2\pi\beta)^{n/2} g(2\pi\beta r)$ , where  $g(x) = x^{-n/2} J_{n/2}(x)$ . Then  $F'_0(r) = \sqrt{\frac{n}{S_0}} (2\pi\beta)^{n/2} g'(2\pi\beta r)(2\pi\beta)$ , where  $g'(x) = -x^{-n/2} J_{(n/2)+1}(x)$ , see [1] eqn. 9.1.30.) Then we write the following: (for the last step, recall the fact that  $\lambda \leq 2 \cdot 2\pi\beta e/n$ , which implies  $2\pi\beta/(\lambda\eta e) \geq$

$n/(2\eta e^2)$

$$\begin{aligned}
\int_{1/(\lambda\eta e)}^{\infty} F_0'(r)^2 r^{n-k} dr &= \frac{n}{S_0} \int_{1/(\lambda\eta e)}^{\infty} J_{(n/2)+1}(2\pi\beta r)^2 r^{-k} dr \cdot (2\pi\beta)^2 \\
&= \frac{n}{S_0} \int_{2\pi\beta/(\lambda\eta e)}^{\infty} J_{(n/2)+1}(t)^2 t^{-k} dt \cdot (2\pi\beta)^{k+1} \\
&\leq \frac{n}{S_0} \frac{4}{\pi\sqrt{3}k} \left(\frac{2\eta e^2}{n}\right)^k \cdot (2\pi\beta)^{k+1}.
\end{aligned} \tag{117}$$

We now combine this with the results of section 3, to show a bound on the variance of  $\vec{b}$  perpendicular to  $\vec{\theta}$ , conditioned on  $a \leq \eta$ .

We can simplify equations (30), (37) and (55) as follows:

$$\begin{aligned}
\left| \int_0^\eta I_{Br} I_{B1} I_2 \frac{da}{a^{n+1}} \right| &\leq \frac{1}{4}(n-2) \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-3} dr, \\
0 \leq \int_0^\eta I_{Ar} I_{A1} I_2 \frac{da}{a^{n+1}} &\leq \frac{\pi^2}{2(n-1)} \cdot \frac{1}{\lambda} \int_{1/(\lambda\eta e)}^{\infty} F_0'(r)^2 r^{n-2} dr + \frac{\pi^2}{2(n-1)} \cdot \frac{17}{\lambda} \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-4} dr, \\
0 \leq \int_0^\eta I_{Cr} I_{C1} I_2 \frac{da}{a^{n+1}} &\leq (2n+9 + \frac{10}{n-3})e\lambda \int_{1/(\lambda\eta e)}^{\infty} F_0(r)^2 r^{n-2} dr.
\end{aligned}$$

Plugging in our bounds for the integrals on the right hand side, we get:

$$\begin{aligned}
\left| \int_0^\eta I_{Br} I_{B1} I_2 \frac{da}{a^{n+1}} \right| &\leq \frac{25}{S_0} \frac{(2\eta)^3}{n} (2\pi\beta)^2, \\
0 \leq \int_0^\eta I_{Ar} I_{A1} I_2 \frac{da}{a^{n+1}} &\leq \frac{5}{n-1} \left( \frac{8}{S_0} (2\eta)^2 (2\pi\beta)^2 + \frac{3600}{S_0} \frac{(2\eta)^4}{n^2} (2\pi\beta)^2 \right), \\
0 \leq \int_0^\eta I_{Cr} I_{C1} I_2 \frac{da}{a^{n+1}} &\leq \left( 1 + \frac{5}{n-3} \right) \frac{640}{S_0} \frac{(2\eta)^2}{n} (2\pi\beta)^2.
\end{aligned}$$

Substituting into equation (26), and using (113), we get:

$$\begin{aligned}
E(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) &\leq \frac{5.20}{\eta} \left( 1 + \frac{1}{n-1} \right) \beta^2 \left( 640(2\eta)^2 + 40(2\eta)^2 + \frac{3200}{n-3} (2\eta)^2 + 50(2\eta)^3 + \frac{18000}{n^2} (2\eta)^4 \right) \\
&\leq (1 + O(\frac{1}{n})) \beta^2 \eta \left( 3536 + 260(2\eta) + O(\frac{1}{n}) \right) \cdot 4.
\end{aligned} \tag{118}$$

Using our assumption that  $2\eta \leq 1/e^2$ , we can rewrite this as

$$\boxed{E(\vec{b}^T (I - \vec{\theta}\vec{\theta}^T) \vec{b} \mid a \leq \eta) \leq \eta \beta^2 (14300 + O(\frac{1}{n}))}. \tag{119}$$

### C.5 The variance of $\vec{b}$ parallel to $\vec{\theta}$

We also get a bound on the variance of  $\vec{b}$  parallel to  $\vec{\theta}$ , conditioned on  $a \leq \eta$ .

Substituting into equations (76), (82) and (93), we get:

$$\begin{aligned} \left| \int_0^\eta K_{Br} K_{B1} K_2 \frac{da}{a^{n+1}} \right| &\leq \frac{25}{S_0} (2\eta)^3 (2\pi\beta)^2, \\ 0 \leq \int_0^\eta K_{Ar} K_{A1} K_2 \frac{da}{a^{n+1}} &\leq \frac{12}{S_0} (2\eta) (2\pi\beta)^2 + \frac{5300}{S_0} \frac{(2\eta)^3}{n^2} (2\pi\beta)^2, \\ 0 \leq \int_0^\eta K_{Cr} K_{C1} K_2 \frac{da}{a^{n+1}} &\leq \frac{700}{S_0} \frac{(2\eta)^3}{n} (2\pi\beta)^2 + \frac{500}{S_0} (2\eta)^3 \left(1 + \frac{5}{n-3}\right) (2\pi\beta)^2. \end{aligned}$$

Substituting into equation (75), and using (113), we get:

$$\begin{aligned} E((\vec{b} \cdot \vec{\theta})^2 | a \leq \eta) &\leq \frac{5.20}{\eta} \left(1 + \frac{1}{n-1}\right) \beta^2 \left(12(2\eta) + 100(2\eta)^3 \left(\frac{53}{n^2} + \frac{1}{2} + \frac{7}{n} + 5 + \frac{25}{n-3}\right)\right) \\ &\leq (1 + O(\frac{1}{n})) \beta^2 \left(63 + 3120(2\eta)^2 + O(\frac{1}{n})\right) \cdot 2. \end{aligned} \tag{120}$$

Using our assumption that  $2\eta \leq 1/e^2$ , we can rewrite this as

$$\boxed{E((\vec{b} \cdot \vec{\theta})^2 | a \leq \eta) \leq \beta^2 (242 + O(\frac{1}{n}))}. \tag{121}$$

## D Spherical Shells

### D.1 Spherical shell with square cross-section

In this section, we give a heuristic analysis of the spherical shell with square cross-section.

We can write  $f = h - g$ , where  $h$  is  $C$  times the indicator function of a ball of radius  $\beta + \delta$  around the origin, and  $g$  is  $C$  times the indicator function of a ball of radius  $\beta$  around the origin. Then its Fourier transform is  $\hat{f} = \hat{h} - \hat{g}$ , where  $\hat{h}$  and  $\hat{g}$  are calculated as in Section 4. We can write this as:  $\hat{f}(\vec{k}) = F_0(|\vec{k}|)$ ,

$$F_0(\rho) = \frac{C}{\rho^{n/2}} (\beta + \delta)^{n/2} J_{n/2}(2\pi(\beta + \delta)\rho) - \frac{C}{\rho^{n/2}} \beta^{n/2} J_{n/2}(2\pi\beta\rho). \tag{122}$$

We are interested in the case where  $\delta \ll \beta$ . Note that interference between the two Bessel functions begins to play a major role when  $\rho \gtrsim 1/(2\pi\delta)$ . We claim that  $F_0(\rho)$  decays quite slowly, out to distance  $\rho \sim 1/(2\pi\delta)$ .

It will be convenient to define

$$K(\beta) = \beta^{n/2} J_{n/2}(2\pi\beta\rho), \tag{123}$$

so we have

$$F_0(\rho) = \frac{C}{\rho^{n/2}} (K(\beta + \delta) - K(\beta)). \tag{124}$$

We can approximate  $F_0(\rho)$  by a simpler expression. First, when  $\delta \lesssim \beta/n$ , we can write  $C$  as follows:

$$C \approx \frac{1}{\sqrt{n\beta^{n-1}\delta B_0}} = \sqrt{\frac{\beta}{n\delta}} \frac{1}{\sqrt{\beta^n B_0}}. \tag{125}$$

Also, when  $\delta$  is sufficiently small (we will elaborate on this point later),

$$F_0(\rho) \approx \frac{C}{\rho^{n/2}} \delta K'(\beta). \quad (126)$$

A straightforward calculation (see [1], equation 9.1.30) shows that

$$K'(\beta) = (2\pi\rho)\beta^{n/2} J_{(n/2)-1}(2\pi\beta\rho). \quad (127)$$

Note that  $K'(\beta)$  is roughly  $2\pi\rho$  times larger than  $K(\beta)$ , so we expect the approximation to be accurate when  $\delta \lesssim 1/(2\pi\rho)$ , or equivalently, when  $\rho \lesssim 1/(2\pi\delta)$ . Combining the above equations, we get the following approximation for  $F_0(\rho)$ :

$$\begin{aligned} F_0(\rho) &\approx \sqrt{\frac{\beta}{n\delta}} \frac{1}{\sqrt{\beta^n B_0}} \cdot \frac{1}{\rho^{n/2}} \delta \cdot (2\pi\rho)\beta^{n/2} J_{(n/2)-1}(2\pi\beta\rho) \\ &= \sqrt{\frac{\beta\delta}{S_0}} \frac{2\pi}{\rho^{(n/2)-1}} J_{(n/2)-1}(2\pi\beta\rho) \text{ when } \rho \lesssim 1/(2\pi\delta), \end{aligned} \quad (128)$$

where  $S_0$  is the surface area of the sphere in  $\mathbb{R}^n$  (note that  $B_0 = S_0/n$ ).

Compared to the Fourier transform of the ball (Section 4), this function decays more slowly, out to distance  $\rho \sim 1/(2\pi\delta)$ . Thus, when we apply the curvelet transform, with significant probability, we can observe fine-scale elements  $a \leq \eta$ , where  $\eta$  shrinks proportional to  $\delta$ . This suggests that a very thin spherical shell (i.e.,  $\delta$  very small) allows us to find the center with very high precision, proportional to  $\delta$ .

## D.2 Spherical shell with Gaussian cross-section

In the next few sections, we will prove Theorem 4, for a spherical shell with Gaussian cross-section.

We begin by proving upper and lower bounds on the normalization factor  $C_f$ . The following identity will be useful: (this follows from the definition of  $F_0(\rho)$  and a change of variables)

$$\int_{\alpha}^{\alpha'} F_0(\rho)^2 \rho^{n-1} d\rho = C_f^2 \varepsilon^n \beta^{2n-2} \int_{2\pi\beta\alpha}^{2\pi\beta\alpha'} \exp(-\frac{1}{2\pi}\varepsilon^2 t^2) J_{(n/2)-1}(t)^2 t dt. \quad (129)$$

Then the  $L^2$  norm of  $\hat{f}(\vec{k})$  is given by:

$$\begin{aligned} \int_{\mathbb{R}^n} |\hat{f}(\vec{k})|^2 d\vec{k} &= S_0 \int_0^\infty F_0(\rho)^2 \rho^{n-1} d\rho \\ &= S_0 C_f^2 \varepsilon^n \beta^{2n-2} \int_0^\infty \exp(-\frac{1}{2\pi}\varepsilon^2 t^2) J_{(n/2)-1}(t)^2 t dt \\ &= S_0 C_f^2 \varepsilon^n \beta^{2n-2} (N_1 + N_2), \end{aligned} \quad (130)$$

where we split the integral into two parts,

$$N_1 = \int_0^{n-2} \exp(-\frac{1}{2\pi}\varepsilon^2 t^2) J_{(n/2)-1}(t)^2 t dt, \quad (131)$$

$$N_2 = \int_{n-2}^\infty \exp(-\frac{1}{2\pi}\varepsilon^2 t^2) J_{(n/2)-1}(t)^2 t dt. \quad (132)$$



We now prove upper bounds on  $N_1$  and  $N_2$ . For  $N_1$ , using trivial upper bounds on  $\exp(-x^2)$  and  $J_\nu(x)^2$  (see [1], eqn. 9.1.60), we get:

$$N_1 \leq \int_0^{n-2} \frac{1}{2} t dt = \frac{1}{4} (n-2)^2. \quad (133)$$

For  $N_2$ , using the upper bound on  $J_\nu(x)^2$  from equation (102), we get:

$$\begin{aligned} N_2 &\leq \int_{n-2}^{\infty} \exp\left(-\frac{1}{2\pi}\varepsilon^2 t^2\right) \frac{4}{\pi\sqrt{3}} dt \\ &= \frac{4}{\pi\sqrt{3}} \sqrt{2\pi} \frac{1}{\varepsilon} \int_{\frac{1}{\sqrt{2\pi}}\varepsilon(n-2)}^{\infty} \exp(-\tau^2) d\tau \\ &\leq \frac{4}{\pi\sqrt{3}} \sqrt{2\pi} \frac{1}{\varepsilon} \frac{\sqrt{\pi}}{2} = 2\sqrt{\frac{2}{3}} \frac{1}{\varepsilon}. \end{aligned} \quad (134)$$

Substituting into (130), we get

$$\int_{\mathbb{R}^n} |\hat{f}(\vec{k})|^2 d\vec{k} \leq S_0 C_f^2 \varepsilon^n \beta^{2n-2} \left(\frac{1}{4}(n-2)^2 + 2\sqrt{\frac{2}{3}} \frac{1}{\varepsilon}\right).$$

We assumed  $\varepsilon \leq \frac{6}{(n-2)^2}$ , and it is easy to check that this implies  $\frac{1}{4}(n-2)^2 \leq 2\sqrt{\frac{2}{3}} \frac{1}{\varepsilon}$ . Thus

$$\int_{\mathbb{R}^n} |\hat{f}(\vec{k})|^2 d\vec{k} \leq S_0 C_f^2 \varepsilon^n \beta^{2n-2} \cdot 4\sqrt{\frac{2}{3}} \frac{1}{\varepsilon}.$$

Setting the left side equal to 1 implies a lower bound on  $C_f^2$ :

$$\boxed{C_f^2 \geq \frac{1}{S_0 \varepsilon^{n-1} \beta^{2n-2}} \cdot \frac{1}{4} \sqrt{\frac{3}{2}}.} \quad (135)$$

Next we show lower bounds on  $N_1$  and  $N_2$ . For  $N_1$  we have a trivial lower bound,

$$N_1 \geq 0. \quad (136)$$

For  $N_2$ , we use the lower bound  $J_\nu(x)^2 > \frac{2}{\pi x} \cos^2(\theta_\nu(x))$  (see equations (103) and (101)). For convenience, we define  $\theta(x) = \theta_\nu(x)$ , suppressing the  $\nu$  subscript. We get:

$$N_2 \geq \frac{2}{\pi} \int_{n-2}^{\infty} \exp\left(-\frac{1}{2\pi}\varepsilon^2 t^2\right) \cos^2(\theta(t)) dt. \quad (137)$$

Note that  $\theta(t)$  is a monotone increasing function (equation (106)), hence it is one-to-one and has a well-defined inverse. We make a change of variables,  $\tau = \theta(t)$ ,  $t = \theta^{-1}(\tau)$ :

$$N_2 \geq \frac{2}{\pi} \int_{\theta(n-2)}^{\infty} \exp\left(-\frac{1}{2\pi}\varepsilon^2 (\theta^{-1}(\tau))^2\right) \cos^2 \tau \cdot (\theta^{-1})'(\tau) d\tau. \quad (138)$$

Note that, whenever  $\tau = \theta(t)$ , we have  $(\theta^{-1})'(\tau) = \frac{1}{\theta'(t)}$ . Hence, by equation (106),

$$1 < (\theta^{-1})'(\tau) < \frac{2}{\sqrt{3}}, \quad \text{for } \tau \geq \theta(n-2). \quad (139)$$

Also note that

$$\begin{aligned}\theta^{-1}(\tau) &= \theta^{-1}(\theta(n-2)) + \int_{\theta(n-2)}^{\tau} (\theta^{-1})'(x) dx \\ &\leq n-2 + \frac{2}{\sqrt{3}}(\tau - \theta(n-2)).\end{aligned}\tag{140}$$

Substituting in, we get:

$$N_2 \geq \frac{2}{\pi} \int_{\theta(n-2)}^{\infty} \exp(-\frac{1}{2\pi}\varepsilon^2 \cdot (n-2 + \frac{2}{\sqrt{3}}(\tau - \theta(n-2)))^2) \cos^2 \tau d\tau.\tag{141}$$

We will use the following simple fact: if a function  $f$  is nonnegative and monotone decreasing on the interval  $[\alpha, \infty)$ , then

$$\int_{\alpha}^{\infty} f(x) \cos^2 x dx \geq \frac{1}{2} \int_{\alpha+\pi}^{\infty} f(x) dx.\tag{142}$$

This follows because

$$\begin{aligned}\int_{\alpha}^{\infty} f(x) \cos^2 x dx &= \sum_{k=0}^{\infty} \int_{\alpha+k\pi}^{\alpha+(k+1)\pi} f(x) \cos^2 x dx \\ &\geq \sum_{k=0}^{\infty} f(\alpha + (k+1)\pi) \int_{\alpha+k\pi}^{\alpha+(k+1)\pi} \cos^2 x dx \\ &= \sum_{k=0}^{\infty} f(\alpha + (k+1)\pi) \frac{\pi}{2} \\ &= \sum_{k=0}^{\infty} f(\alpha + (k+1)\pi) \frac{1}{2} \int_{\alpha+(k+1)\pi}^{\alpha+(k+2)\pi} dx \\ &\geq \sum_{k=0}^{\infty} \frac{1}{2} \int_{\alpha+(k+1)\pi}^{\alpha+(k+2)\pi} f(x) dx \\ &= \frac{1}{2} \int_{\alpha+\pi}^{\infty} f(x) dx.\end{aligned}\tag{143}$$

Using the above fact, and a change of variables, we get

$$\begin{aligned}N_2 &\geq \frac{1}{\pi} \int_{\theta(n-2)+\pi}^{\infty} \exp(-\frac{1}{2\pi}\varepsilon^2 \cdot (n-2 + \frac{2}{\sqrt{3}}(\tau - \theta(n-2)))^2) d\tau \\ &= \frac{\sqrt{3}}{2\pi} \int_{n-2+\frac{2\pi}{\sqrt{3}}}^{\infty} \exp(-\frac{1}{2\pi}\varepsilon^2 x^2) dx \\ &= \frac{\sqrt{3}}{\sqrt{2\pi}} \frac{1}{\varepsilon} \int_{\frac{1}{\sqrt{2\pi}}\varepsilon(n-2+\frac{2\pi}{\sqrt{3}})}^{\infty} \exp(-y^2) dy.\end{aligned}\tag{144}$$

Recall that we assumed  $\varepsilon \leq \frac{1}{n+2}$ . This implies  $\frac{1}{\sqrt{2\pi}}\varepsilon(n-2+\frac{2\pi}{\sqrt{3}}) \leq \frac{1}{\sqrt{2\pi}}$ . So, substituting in and integrating numerically, we get that

$$N_2 \geq \frac{\sqrt{3}}{\sqrt{2\pi}} \frac{1}{\varepsilon} \int_{\frac{1}{\sqrt{2\pi}}}^{\infty} \exp(-y^2) dy \geq \frac{1}{4\varepsilon}.\tag{145}$$

Substituting into (130), we get that

$$\int_{\mathbb{R}^n} |\hat{f}(\vec{k})|^2 d\vec{k} \geq S_0 C_f^2 \varepsilon^n \beta^{2n-2} (0 + \frac{1}{4\varepsilon}) = S_0 C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \frac{1}{4}. \quad (146)$$

Setting the left side equal to 1 implies an upper bound on  $C_f^2$ :

$$\boxed{C_f^2 \leq \frac{4}{S_0 \varepsilon^{n-1} \beta^{2n-2}}}. \quad (147)$$

### D.3 The low-frequency components

Next, we show that  $\hat{f}(\vec{k})$  has very little power at low frequencies, corresponding to coarse scales  $a \geq 1$ . This justifies our use of the curvelet transform, which effectively ignores these low-frequency components (recall Theorems 1 and 2).

The total amount of power at frequencies less than  $z$  (for any  $z \geq 0$ ) is given by:

$$\begin{aligned} \int_{|\vec{k}| \leq z} |\hat{f}(\vec{k})|^2 d\vec{k} &= S_0 \int_0^z F_0(\rho)^2 \rho^{n-1} d\rho \\ &= S_0 C_f^2 \varepsilon^n \beta^{2n-2} \int_0^{2\pi\beta z} \exp(-\frac{1}{2\pi} \varepsilon^2 t^2) J_{(n/2)-1}(t)^2 t dt. \end{aligned} \quad (148)$$

(We used equation (129).) Using a trivial upper bound  $\exp(-x^2) \leq 1$ , and the upper bound for  $J_\nu(x)$  (when  $x$  is small) from equation (96), we get:

$$\begin{aligned} \int_{|\vec{k}| \leq z} |\hat{f}(\vec{k})|^2 d\vec{k} &\leq S_0 C_f^2 \varepsilon^n \beta^{2n-2} \int_0^{2\pi\beta z} \frac{1}{\pi(n-2)} \left(\frac{te}{n-2}\right)^{n-2} t dt \\ &= S_0 C_f^2 \varepsilon^n \beta^{2n-2} \cdot \frac{e^{n-2}}{\pi(n-2)^{n-1}} \cdot \frac{1}{n} (2\pi\beta z)^n. \end{aligned} \quad (149)$$

Using our upper bound on  $C_f^2$  (equation (147)), we get

$$\begin{aligned} \int_{|\vec{k}| \leq z} |\hat{f}(\vec{k})|^2 d\vec{k} &\leq 4\varepsilon \cdot \frac{e^{n-2}}{\pi(n-2)^{n-1}} \cdot \frac{1}{n} (2\pi\beta z)^n \\ &\leq \frac{4\varepsilon}{\pi e^2} \cdot \frac{e^n}{(n-2)^n} \cdot (2\pi\beta z)^n. \end{aligned} \quad (150)$$

Now, fix  $z = \frac{n-2}{2\pi\beta e}$ . Recall that  $\lambda \geq \frac{2\pi\beta e}{n-2}$ , hence  $1/\lambda \leq z$ . Then we have

$$\boxed{\int_{|\vec{k}| \leq 1/\lambda} |\hat{f}(\vec{k})|^2 d\vec{k} \leq \frac{4\varepsilon}{\pi e^2} \leq \frac{\varepsilon}{5}}. \quad (151)$$

Recall that we assumed  $\varepsilon \leq \frac{1}{n+2}$ . So the frequencies below  $1/\lambda$  only constitute a small fraction of the total probability mass. This justifies our use of the curvelet transform, with this choice of the parameter  $\lambda$ .

#### D.4 The probability of measuring a fine-scale element

We give a lower bound on the probability of measuring the scale variable to be small,  $a \leq \eta_c$ , where

$$\eta_c = \frac{\delta}{\tilde{\beta}} \frac{(n-2)}{e}. \quad (152)$$

We will show that  $a \leq \eta_c$  with at least constant probability.

First, we write  $\Pr[a \leq \eta_c]$  as follows, using Section 3.1 and equation (129):

$$\begin{aligned} \Pr[a \leq \eta_c] &\geq S_0 \int_{1/(\lambda\eta_c)}^{\infty} F_0(\rho)^2 \rho^{n-1} d\rho \\ &= S_0 C_f^2 \varepsilon^n \beta^{2n-2} \int_{2\pi\beta/(\lambda\eta_c)}^{\infty} \exp(-\frac{1}{2\pi}\varepsilon^2 t^2) J_{(n/2)-1}(t)^2 t dt. \end{aligned} \quad (153)$$

The lower limit of integration can be simplified, by substituting in the definitions of  $\lambda$  and  $\eta_c$ ,

$$\frac{2\pi\beta}{\lambda\eta_c} = 2\pi\beta \left( \frac{2\pi\tilde{\beta}e}{n-2} \right)^{-1} \left( \frac{\delta}{\tilde{\beta}} \frac{(n-2)}{e} \right)^{-1} = 2\pi\beta(2\pi\delta)^{-1} = \frac{\beta}{\delta} = \frac{1}{\varepsilon}. \quad (154)$$

This integral is similar to the integral  $N_2$  which we encountered earlier (equation (132)). We can get a lower bound using the same technique (equations (137) - (144)); in particular, note that  $1/\varepsilon \geq n-2$ , as required in that calculation (this holds because we assumed  $\varepsilon \leq 1/(n+2)$ ). This leads to:

$$\Pr[a \leq \eta_c] \geq S_0 C_f^2 \varepsilon^n \beta^{2n-2} \cdot \frac{\sqrt{3}}{\sqrt{2\pi}} \frac{1}{\varepsilon} \int_{\frac{1}{\sqrt{2\pi}}\varepsilon(\frac{1}{\varepsilon} + \frac{2\pi}{\sqrt{3}})}^{\infty} \exp(-y^2) dy. \quad (155)$$

The lower limit of integration can be written as

$$\frac{1}{\sqrt{2\pi}}\varepsilon \left( \frac{1}{\varepsilon} + \frac{2\pi}{\sqrt{3}} \right) = \frac{1}{\sqrt{2\pi}} \left( 1 + \frac{2\pi}{\sqrt{3}}\varepsilon \right) \leq \frac{1}{\sqrt{2\pi}} \cdot 2 = \sqrt{\frac{2}{\pi}}, \quad (156)$$

where the last inequality follows because  $\varepsilon \leq \frac{1}{5}$ . Thus we can lower-bound our integral as follows:

$$\begin{aligned} \Pr[a \leq \eta_c] &\geq S_0 C_f^2 \varepsilon^n \beta^{2n-2} \cdot \frac{\sqrt{3}}{\sqrt{2\pi}} \frac{1}{\varepsilon} \int_{\sqrt{2/\pi}}^{\infty} \exp(-y^2) dy \\ &> S_0 C_f^2 \varepsilon^n \beta^{2n-2} \cdot \frac{0.15}{\varepsilon} \\ &= (0.15) S_0 C_f^2 \varepsilon^{n-1} \beta^{2n-2}. \end{aligned} \quad (157)$$

Now, using our lower bound for  $C_f^2$  (equation (135)), we get:

$\Pr[a \leq \eta_c] \geq (0.15) \cdot \frac{1}{4} \sqrt{\frac{3}{2}} > 0.045.$

(158)

#### D.5 Some more integrals

Our next goal is to bound the variance of  $\vec{b}$ . We begin by proving upper bounds on certain integrals involving  $F_0(r)$  and  $F'_0(r)$ . Then, in the following two sections, we will bound the variance of  $\vec{b}$  in the directions orthogonal and parallel to  $\vec{\theta}$ .

First, we consider integrals of the following form, where  $k \geq 1$ :

$$\int_{\alpha}^{\infty} F_0(r)^2 r^{n-k} dr. \quad (159)$$

Using the definition of  $F_0(r)$ , and a change of variables,

$$\int_{\alpha}^{\infty} F_0(r)^2 r^{n-k} dr = C_f^2 \varepsilon^n (2\pi)^{k-1} \beta^{2n+k-3} \int_{2\pi\beta\alpha}^{\infty} \exp(-\frac{1}{2\pi} \varepsilon^2 t^2) J_{(n/2)-1}(t)^2 \frac{dt}{t^{k-2}}.$$

We will upper-bound this integral, assuming that  $\alpha \geq \frac{n-2}{2\pi\beta}$ . First, using equation (102), we have  $J_{(n/2)-1}(t)^2 \leq \frac{4}{\pi\sqrt{3}} \frac{1}{t}$ , and we get:

$$\int_{\alpha}^{\infty} F_0(r)^2 r^{n-k} dr \leq C_f^2 \varepsilon^n (2\pi)^{k-1} \beta^{2n+k-3} \cdot \frac{4}{\pi\sqrt{3}} \int_{2\pi\beta\alpha}^{\infty} \exp(-\frac{1}{2\pi} \varepsilon^2 t^2) \frac{dt}{t^{k-1}}.$$

Next, we use a simple inequality:  $t^{k-1} \geq (2\pi\beta\alpha)^{k-1}$ , whenever  $t \geq 2\pi\beta\alpha$ . Thus,

$$\int_{\alpha}^{\infty} F_0(r)^2 r^{n-k} dr \leq C_f^2 \varepsilon^n \beta^{2n-2} \cdot \frac{4}{\pi\sqrt{3}} \frac{1}{\alpha^{k-1}} \int_{2\pi\beta\alpha}^{\infty} \exp(-\frac{1}{2\pi} \varepsilon^2 t^2) dt.$$

The integral on the right hand side can be bounded as follows:

$$\int_{2\pi\beta\alpha}^{\infty} \exp(-\frac{1}{2\pi} \varepsilon^2 t^2) dt = \int_{\frac{1}{\sqrt{2\pi}} \varepsilon \cdot 2\pi\beta\alpha}^{\infty} \exp(-\tau^2) d\tau \cdot \sqrt{2\pi} \frac{1}{\varepsilon} \leq \frac{\sqrt{\pi}}{2} \cdot \sqrt{2\pi} \frac{1}{\varepsilon} = \frac{\pi}{\sqrt{2}} \frac{1}{\varepsilon}.$$

Substituting in, we get:

$$\boxed{\int_{\alpha}^{\infty} F_0(r)^2 r^{n-k} dr \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \frac{4}{\sqrt{6}} \frac{1}{\alpha^{k-1}}, \quad \text{for } \alpha \geq \frac{n-2}{2\pi\beta}.} \quad (160)$$

Next, we consider integrals of the following form, where  $k = 1, 2$ :

$$\int_{\alpha}^{\infty} F'_0(r)^2 r^{n-k} dr. \quad (161)$$

In order to calculate  $F'_0(r)$ , we write  $F_0(r)$  in the following form:

$$F_0(r) = C_f \varepsilon^{n/2} (2\pi) \beta^n \cdot (2\pi\beta)^{(n/2)-1} \cdot H(r) \cdot K(2\pi\beta r),$$

where

$$H(x) = \exp(-\pi \varepsilon^2 \beta^2 x^2), \quad K(x) = \frac{1}{x^{(n/2)-1}} J_{(n/2)-1}(x).$$

Then  $F'_0(r)$  can be written as

$$F'_0(r) = F_0^{(a)}(r) + F_0^{(b)}(r),$$

where

$$\begin{aligned} F_0^{(a)}(r) &= C_f \varepsilon^{n/2} (2\pi) \beta^n \cdot (2\pi\beta)^{(n/2)-1} \cdot H'(r) \cdot K(2\pi\beta r), \\ F_0^{(b)}(r) &= C_f \varepsilon^{n/2} (2\pi) \beta^n \cdot (2\pi\beta)^{(n/2)-1} \cdot H(r) \cdot K'(2\pi\beta r) (2\pi\beta). \end{aligned}$$

We can expand this out. Note that (see [1], equation (9.1.30))

$$H'(x) = \exp(-\pi\varepsilon^2\beta^2x^2)(-2\pi\varepsilon^2\beta^2x), \quad K'(x) = -\frac{1}{x^{(n/2)-1}}J_{n/2}(x).$$

Substituting in, we get

$$\begin{aligned} F_0^{(a)}(r) &= C_f \varepsilon^{n/2} (2\pi) \beta^n \frac{1}{r^{(n/2)-1}} \cdot \exp(-\pi\varepsilon^2\beta^2r^2)(-2\pi\varepsilon^2\beta^2r) J_{(n/2)-1}(2\pi\beta r), \\ F_0^{(b)}(r) &= C_f \varepsilon^{n/2} (2\pi) \beta^n \frac{1}{r^{(n/2)-1}} \cdot \exp(-\pi\varepsilon^2\beta^2r^2)(-1) J_{n/2}(2\pi\beta r) (2\pi\beta). \end{aligned}$$

Thus  $F'_0(r)$  is given by:

$$F'_0(r) = C_f \varepsilon^{n/2} (2\pi) \beta^n \frac{1}{r^{(n/2)-1}} \exp(-\pi\varepsilon^2\beta^2r^2) \cdot \left( -2\pi\varepsilon^2\beta^2r J_{(n/2)-1}(2\pi\beta r) - 2\pi\beta J_{n/2}(2\pi\beta r) \right). \quad (162)$$

Substituting into our integral, and performing a change of variables, we get:

$$\int_{\alpha}^{\infty} F'_0(r)^2 r^{n-k} dr = C_f^2 \varepsilon^n (2\pi)^{k-1} \beta^{2n+k-3} \int_{2\pi\beta\alpha}^{\infty} \exp(-\frac{1}{2\pi}\varepsilon^2t^2) \left( \varepsilon^2\beta t J_{(n/2)-1}(t) + 2\pi\beta J_{n/2}(t) \right)^2 \frac{dt}{t^{k-2}}. \quad (163)$$

We will upper-bound this integral, assuming that  $\alpha \geq \frac{n}{2\pi\beta}$ . First, using equation (102), we have  $J_{n/2}(t)^2 \leq \frac{4}{\pi\sqrt{3}} \frac{1}{t}$ , and similarly for  $J_{(n/2)-1}(t)^2$ . So we get:

$$\int_{\alpha}^{\infty} F'_0(r)^2 r^{n-k} dr \leq C_f^2 \varepsilon^n (2\pi)^{k-1} \beta^{2n+k-3} \int_{2\pi\beta\alpha}^{\infty} \exp(-\frac{1}{2\pi}\varepsilon^2t^2) \left( \varepsilon^2\beta t + 2\pi\beta \right)^2 \left( \frac{4}{\pi\sqrt{3}} \frac{1}{t} \right) \frac{dt}{t^{k-2}}. \quad (164)$$

Changing variables and rearranging, we get:

$$\begin{aligned} &\int_{\alpha}^{\infty} F'_0(r)^2 r^{n-k} dr \\ &\leq C_f^2 \varepsilon^n (2\pi)^{k-1} \beta^{2n+k-3} \int_{\sqrt{2\pi\varepsilon\beta\alpha}}^{\infty} \exp(-\tau^2) (\sqrt{2\pi\varepsilon\beta\alpha} \tau + 2\pi\beta)^2 \frac{4}{\pi\sqrt{3}} \frac{d\tau}{\tau^{k-1}} \cdot \left( \frac{1}{\sqrt{2\pi}} \varepsilon \right)^{k-2} \\ &= C_f^2 \varepsilon^n (2\pi)^{k-1} \beta^{2n+k-3} \cdot \frac{4}{\pi\sqrt{3}} \left( \frac{1}{\sqrt{2\pi}} \varepsilon \right)^{k-2} \int_{\sqrt{2\pi\varepsilon\beta\alpha}}^{\infty} \exp(-\tau^2) (2\pi\varepsilon^2\beta^2\tau^2 + 2(2\pi)^{3/2}\varepsilon\beta^2\tau + (2\pi)^2\beta^2) \frac{d\tau}{\tau^{k-1}} \\ &= C_f^2 \varepsilon^{n+k-2} \sqrt{2\pi}^k \beta^{2n+k-1} \cdot \frac{4}{\pi\sqrt{3}} \int_{\sqrt{2\pi\varepsilon\beta\alpha}}^{\infty} \exp(-\tau^2) (2\pi\varepsilon^2\tau^2 + 2(2\pi)^{3/2}\varepsilon\tau + (2\pi)^2) \frac{d\tau}{\tau^{k-1}}. \end{aligned} \quad (165)$$

We can handle integrals of the form

$$\int_a^{\infty} \exp(-\tau^2) \frac{d\tau}{\tau^\ell} \quad (166)$$

as follows. When  $\ell \geq 0$ , we have

$$\int_a^{\infty} \exp(-\tau^2) \frac{d\tau}{\tau^\ell} \leq \frac{1}{a^\ell} \int_a^{\infty} \exp(-\tau^2) d\tau \leq \frac{1}{a^\ell} \cdot \frac{\sqrt{\pi}}{2}. \quad (167)$$

When  $\ell = -1$ , we have

$$\int_a^{\infty} \exp(-\tau^2) \tau d\tau \leq \int_0^{\infty} \exp(-\tau^2) \tau d\tau = -\frac{1}{2} \exp(-\tau^2) \Big|_0^{\infty} = \frac{1}{2}. \quad (168)$$

When  $\ell = -2$ , we have

$$\int_a^\infty \exp(-\tau^2) \tau^2 d\tau \leq \int_0^\infty \exp(-\tau^2) \tau^2 d\tau = -\frac{1}{2} \exp(-\tau^2) \tau \Big|_0^\infty - \int_0^\infty -\frac{1}{2} \exp(-\tau^2) d\tau = \frac{\sqrt{\pi}}{4} < 0.45. \quad (169)$$

Now, we can upper-bound our integral in the  $k = 2$  case:

$$\begin{aligned} & \int_\alpha^\infty F'_0(r)^2 r^{n-2} dr \\ & \leq C_f^2 \varepsilon^n (2\pi) \beta^{2n+1} \cdot \frac{4}{\pi\sqrt{3}} \int_{\sqrt{2\pi\varepsilon\beta\alpha}}^\infty \exp(-\tau^2) (2\pi\varepsilon^2 \tau^2 + 2(2\pi)^{3/2} \varepsilon \tau + (2\pi)^2) \frac{d\tau}{\tau} \\ & \leq C_f^2 \varepsilon^n (2\pi) \beta^{2n+1} \cdot \frac{4}{\pi\sqrt{3}} \left( 2\pi\varepsilon^2 \cdot \frac{1}{2} + 2(2\pi)^{3/2} \varepsilon \cdot \frac{\sqrt{\pi}}{2} + (2\pi)^2 \cdot \frac{1}{\sqrt{2\pi\varepsilon\beta\alpha}} \frac{\sqrt{\pi}}{2} \right) \\ & = C_f^2 \varepsilon^n (2\pi) \beta^{2n+1} \cdot \frac{8}{\sqrt{3}} \left( \frac{1}{2} \varepsilon^2 + \sqrt{2\pi} \varepsilon + \frac{\pi}{\sqrt{2}} \frac{1}{\varepsilon\beta\alpha} \right), \quad \text{for } \alpha \geq \frac{n}{2\pi\beta}. \end{aligned} \quad (170)$$

We can also upper-bound our integral in the  $k = 1$  case:

$$\begin{aligned} & \int_\alpha^\infty F'_0(r)^2 r^{n-1} dr \\ & \leq C_f^2 \varepsilon^{n-1} \sqrt{2\pi} \beta^{2n} \cdot \frac{4}{\pi\sqrt{3}} \int_{\sqrt{2\pi\varepsilon\beta\alpha}}^\infty \exp(-\tau^2) (2\pi\varepsilon^2 \tau^2 + 2(2\pi)^{3/2} \varepsilon \tau + (2\pi)^2) d\tau \\ & \leq C_f^2 \varepsilon^{n-1} \sqrt{2\pi} \beta^{2n} \cdot \frac{4}{\pi\sqrt{3}} \left( 2\pi\varepsilon^2 \cdot \frac{\sqrt{\pi}}{4} + 2(2\pi)^{3/2} \varepsilon \cdot \frac{1}{2} + (2\pi)^2 \cdot \frac{\sqrt{\pi}}{2} \right) \\ & = C_f^2 \varepsilon^{n-1} \sqrt{2\pi} \beta^{2n} \cdot \frac{8}{\sqrt{3}} \left( \frac{\sqrt{\pi}}{4} \varepsilon^2 + \sqrt{2\pi} \varepsilon + \pi^{3/2} \right), \quad \text{for } \alpha \geq \frac{n}{2\pi\beta}. \end{aligned} \quad (171)$$

## D.6 The variance of $\vec{b}$ orthogonal to $\vec{\theta}$

We now bound the variance of  $\vec{b}$  orthogonal to  $\vec{\theta}$ , conditioned on observing  $a \leq \eta_c$ . Recall that  $\eta_c = \frac{\delta}{\beta} \frac{(n-2)}{\varepsilon}$ .

We start with the results of Section 3. From equation (30), we get:

$$\left| \int_0^\eta I_{Br} I_{B1} I_2 \frac{da}{a^{n+1}} \right| \leq \frac{1}{4} (n-2) \lambda \eta e \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-2} dr. \quad (172)$$

(We used the fact that  $1 \leq r \cdot \lambda \eta e$ , for all  $r$  in this interval.) From equations (37), (39) and (42), we get:

$$\begin{aligned} 0 \leq \int_0^\eta I_{Ar} I_{A1} I_2 \frac{da}{a^{n+1}} & \leq \frac{\pi^2}{2(n-1)} \cdot \frac{1}{\lambda} \int_{1/(\lambda\eta e)}^\infty F'_0(r)^2 r^{n-2} dr + \\ & \frac{\pi^2}{2(n-1)} \cdot 24e^2 \lambda \eta^2 \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-2} dr. \end{aligned} \quad (173)$$

(Again, we used the fact that  $1 \leq r \cdot \lambda \eta e$ , for all  $r$  in this interval.) From equation (55), we get:

$$0 \leq \int_0^\eta I_{Cr} I_{C1} I_2 \frac{da}{a^{n+1}} \leq 2(n+10) e \lambda \int_{1/(\lambda\eta e)}^\infty F_0(r)^2 r^{n-2} dr. \quad (174)$$

(We used the fact that  $\frac{8}{3} + 2(n-2)(1 + \frac{5}{n-3}) = \frac{8}{3} + 2(n+3 + \frac{5}{n-3}) \leq 2(n+10)$ , assuming  $n \geq 4$ .)

Now we fix  $\eta = \eta_c$ , and we upper-bound these integrals, using equations (160) and (170). (Note that, in order to apply these results, we must have  $\frac{1}{\lambda\eta_c e} \geq \frac{n}{2\pi\beta}$ . Recall that  $\lambda = \frac{2\pi\tilde{\beta}e}{n-2}$ , and  $\eta_c = \frac{\delta}{\beta} \frac{(n-2)}{e}$ . Also, recall that we assumed  $\varepsilon \leq \frac{1}{en}$ . Then  $\lambda\eta_c = 2\pi\delta = 2\pi\varepsilon\beta$ , and  $\frac{1}{\lambda\eta_c e} = \frac{1}{2\pi\beta\varepsilon e} \geq \frac{n}{2\pi\beta}$ , as desired.) After some tedious calculation, we get:

$$\left| \int_0^{\eta_c} I_{Br} I_{B1} I_2 \frac{da}{a^{n+1}} \right| \leq (n-2) \cdot C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot 120 \varepsilon^2 \beta^2, \quad (175)$$

$$0 \leq \int_0^{\eta_c} I_{Ar} I_{A1} I_2 \frac{da}{a^{n+1}} \leq 5 \cdot C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \varepsilon \beta^2 (12000 \varepsilon^2 + 9 \varepsilon + 80), \quad (176)$$

$$0 \leq \int_0^{\eta_c} I_{Cr} I_{C1} I_2 \frac{da}{a^{n+1}} \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot 2600 \varepsilon \beta^2 \left(1 + \frac{12}{n-2}\right) \cdot S. \quad (177)$$

Then, by combining these equations and using our assumption that  $\varepsilon \leq O(\frac{1}{n^2})$ , we get:

$$\begin{aligned} & \int_0^{\eta_c} \left( I_{Ar} I_{A1} I_2 + 2 I_{Br} I_{B1} I_2 + I_{Cr} I_{C1} I_2 \right) \frac{da}{a^{n+1}} \\ & \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \varepsilon \beta^2 (60000 \varepsilon^2 + 45 \varepsilon + 400 + 2(n-2)120 \varepsilon + 2600(1 + \frac{12}{n-2})S) \\ & \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \varepsilon \beta^2 (400 + 2600 + O(\frac{1}{n})) \cdot S. \end{aligned} \quad (178)$$

Next, recall from equation (157) that:

$$\Pr[a \leq \eta_c] \geq (0.15) S_0 C_f^2 \varepsilon^{n-1} \beta^{2n-2}. \quad (179)$$

Finally, by substituting into equation (26) and simplifying, we get a bound on the variance of  $\vec{b}$ , in the subspace orthogonal to  $\vec{\theta}$ , conditioned on observing  $a \leq \eta_c$ :

$$\boxed{\mathbb{E}(\vec{b}^T (I - \vec{\theta} \vec{\theta}^T) \vec{b} \mid a \leq \eta_c) \leq (n-1) \varepsilon \beta^2 (507 + O(\frac{1}{n})) \cdot S.} \quad (180)$$

## D.7 The variance of $\vec{b}$ parallel to $\vec{\theta}$

We now bound the variance of  $\vec{b}$  parallel to  $\vec{\theta}$ , conditioned on observing  $a \leq \eta_c$ . Recall that  $\eta_c = \frac{\delta}{\beta} \frac{(n-2)}{e}$ .

We start with the results of Section 3. From equation (76), we get:

$$\left| \int_0^\eta K_{Br} K_{B1} K_2 \frac{da}{a^{n+1}} \right| \leq \frac{1}{4} (n-1)(n-2) \int_{1/(\lambda\eta_e)}^\infty F_0(r)^2 r^{n-3} dr. \quad (181)$$

From equations (82), (83) and (85), we get:

$$\begin{aligned} 0 \leq \int_0^\eta K_{Ar} K_{A1} K_2 \frac{da}{a^{n+1}} & \leq 2 \int_{1/(\lambda\eta_e)}^\infty F'_0(r)^2 r^{n-1} dr + \\ & 2 \cdot \frac{8}{3} \pi^2 \int_{1/(\lambda\eta_e)}^\infty F_0(r)^2 r^{n-3} dr. \end{aligned} \quad (182)$$

From equation (93), we get:

$$0 \leq \int_0^\eta K_{Cr} K_{C1} K_2 \frac{da}{a^{n+1}} \leq \frac{\pi^2}{2} (n^2 + 3n + 3) \int_{1/(\lambda\eta_e)}^\infty F_0(r)^2 r^{n-3} dr. \quad (183)$$



(We used the fact that  $\frac{2\pi^2}{3}(n-1) + \frac{\pi^2}{2}(n-2)^2(1 + \frac{5}{n-3}) = \frac{\pi^2}{2}(n^2 + \frac{7}{3}n - \frac{7}{3} + \frac{5}{n-3}) \leq \frac{\pi^2}{2}(n^2 + 3n + 3)$ , assuming  $n \geq 4$ .)

Now we fix  $\eta = \eta_c$ , and we upper-bound these integrals, using equations (160) and (171). (Note that, in order to apply these results, we must have  $\frac{1}{\lambda\eta_c e} \geq \frac{n}{2\pi\beta}$ . Recall that  $\lambda = \frac{2\pi\beta e}{n-2}$ , and  $\eta_c = \frac{\delta}{\beta} \frac{(n-2)}{e}$ . Also, recall that we assumed  $\varepsilon \leq \frac{1}{en}$ . Then  $\lambda\eta_c = 2\pi\delta = 2\pi\varepsilon\beta$ , and  $\frac{1}{\lambda\eta_c e} = \frac{1}{2\pi\beta\varepsilon} \geq \frac{n}{2\pi\beta}$ , as desired.)

After some tedious calculation, we get:

$$\left| \int_0^{\eta_c} K_{Br} K_{B1} K_2 \frac{da}{a^{n+1}} \right| \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot n^2 \cdot 120 \varepsilon^2 \beta^2, \quad (184)$$

$$0 \leq \int_0^{\eta_c} K_{Ar} K_{A1} K_2 \frac{da}{a^{n+1}} \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \beta^2 (25200 \varepsilon^2 + 60 \varepsilon + 132), \quad (185)$$

$$0 \leq \int_0^{\eta_c} K_{Cr} K_{C1} K_2 \frac{da}{a^{n+1}} \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot 2400(n^2 + 3n + 3) \varepsilon^2 \beta^2. \quad (186)$$

Then, by combining these equations and using our assumption that  $\varepsilon \leq O(\frac{1}{n^2})$ , we get:

$$\begin{aligned} & \int_0^{\eta_c} \left( K_{Ar} K_{A1} K_2 - 2K_{Br} K_{B1} K_2 + K_{Cr} K_{C1} K_2 \right) \frac{da}{a^{n+1}} \\ & \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \beta^2 (25200 \varepsilon^2 + 60 \varepsilon + 132 + 2 \cdot 120 n^2 \varepsilon^2 + 2400(n^2 + 3n + 3) \varepsilon^2) \\ & \leq C_f^2 \varepsilon^{n-1} \beta^{2n-2} \cdot \beta^2 (132 + O(\frac{1}{n^2})). \end{aligned} \quad (187)$$

Next, recall from equation (157) that:

$$\Pr[a \leq \eta_c] \geq (0.15) S_0 C_f^2 \varepsilon^{n-1} \beta^{2n-2}. \quad (188)$$

Finally, by substituting into equation (75) and simplifying, we get a bound on the variance of  $\vec{b}$ , in the direction  $\vec{\theta}$ , conditioned on observing  $a \leq \eta_c$ :

$$\boxed{\mathbb{E}((\vec{b} \cdot \vec{\theta})^2 \mid a \leq \eta_c) \leq \beta^2 (23 + O(\frac{1}{n^2}))}. \quad (189)$$

## E A Fast Quantum Curvelet Transform

### E.1 The Discrete Curvelet Transform

First, we argue that the discrete Fourier transform approximates the continuous Fourier transform, in the sense described in Section 6.1. This follows from the definitions of the different transforms. Recall the continuous Fourier transform that takes a function on  $\mathbb{R}^n$  to a function on  $\mathbb{R}^n$ :

$$\mathcal{F}_{cont}(f)(\vec{k}) = \int_{\mathbb{R}^n} f(\vec{x}) e^{-2\pi i \vec{k} \cdot \vec{x}} d\vec{x}, \quad (190)$$

$$\mathcal{F}_{cont}^{-1}(g)(\vec{x}) = \int_{\mathbb{R}^n} g(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{x}} d\vec{k}. \quad (191)$$

Now consider the Fourier transform that takes a function on the cube  $C = [-L, L]^n$  (or equivalently, a function on  $\mathbb{R}^n$  that is periodic with respect to the lattice  $(2L\mathbb{Z})^n$ ) to a function on the (dual) lattice  $\hat{C} = (\frac{1}{2L}\mathbb{Z})^n$ . We refer to this as the “semi-discrete” Fourier transform:

$$\mathcal{F}_{semi}(f)(\vec{k}) = (\frac{1}{2L})^{n/2} \int_C f(\vec{x}) e^{-2\pi i \vec{k} \cdot \vec{x}} d\vec{x}, \quad (192)$$

$$\mathcal{F}_{semi}^{-1}(g)(\vec{x}) = (\frac{1}{2L})^{n/2} \sum_{\vec{k} \in \hat{C}} g(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{x}}. \quad (193)$$

Also recall the discrete Fourier transform, that takes a function on  $Z = (\sigma\mathbb{Z})^n \cap [-L, L]^n$  to a function on  $\hat{Z} = (\frac{1}{2L}\mathbb{Z})^n \cap [-\frac{1}{2\sigma}, \frac{1}{2\sigma}]^n$ :

$$\mathcal{F}_{dis}(f)(\vec{k}) = (\frac{\sigma}{2L})^{n/2} \sum_{\vec{x} \in Z} f(\vec{x}) e^{-2\pi i \vec{k} \cdot \vec{x}}, \quad (194)$$

$$\mathcal{F}_{dis}^{-1}(g)(\vec{x}) = (\frac{\sigma}{2L})^{n/2} \sum_{\vec{k} \in \hat{Z}} g(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{x}}. \quad (195)$$

We are given a function  $f_{cont}$  on  $\mathbb{R}^n$  that vanishes outside the cube  $C$ . We define a function  $f_{semi}$  on  $C$  by restriction,  $f_{semi} = f_{cont}|_C$ . Then it follows from the definitions that  $\hat{f}_{semi} = (\frac{1}{2L})^{n/2} \hat{f}_{cont}|_{\hat{C}}$ .

Recall that  $\hat{f}_{cont}$  has most of its probability mass inside the cube  $[-\frac{1}{2\sigma}, \frac{1}{2\sigma}]^n$ . Then the same should be true for  $\hat{f}_{semi}$ . Now define a function  $\hat{f}_{dis}$  on  $\hat{Z}$  by restriction,  $\hat{f}_{dis} = \hat{f}_{semi}|_{\hat{Z}}$ . Then, using the definitions, we see that  $f_{dis} \approx \sigma^{n/2} f_{semi}|_Z$ .

Note that an example of a discrete curvelet transform (over  $\mathbb{R}^2$  or  $\mathbb{R}^3$ ) can be found in [6]. There, the frequency space is partitioned into concentric cubes according to the scale  $a$ , and these are divided into wedges according to the direction  $\vec{\theta}$ . For our purposes, we will use a tiling based on concentric balls (in  $\mathbb{R}^n$ ), which more closely approximates the continuous curvelet transform defined in Section 2.

As a classical computation, the discrete curvelet transform can be implemented using the fast Fourier transform [6] (see in particular the “wrapping” method). We will use these ideas to implement a quantum curvelet transform. The following discussion will be self-contained; but for readers who are familiar with [6], we mention that we omit the “wrapping” step. Our transform produces curvelet coefficients that are somewhat oversampled, but this does not cause any problems in our situation.

## E.2 Constructing the Window Functions $\chi_{a,\vec{\theta}}(\vec{k})$

We will construct two families of window functions  $\chi_{a,\vec{\theta}}(\vec{k})$ , for which the operation  $\mathcal{X}$  can be performed efficiently.

First, suppose we have some partition of the frequency domain into disjoint subsets,  $(\mathbb{Z}_M)^n = \bigcup_{a,\vec{\theta}} S_{a,\vec{\theta}}$ , such that given any point  $\vec{k} \in (\mathbb{Z}_M)^n$ , we can efficiently compute which set  $S_{a,\vec{\theta}}$  contains  $\vec{k}$ . Define the window functions to be the indicator functions for these sets,

$$\chi_{a,\vec{\theta}}(\vec{k}) = \begin{cases} 1, & \text{if } \vec{k} \in S_{a,\vec{\theta}}, \\ 0, & \text{otherwise.} \end{cases} \quad (196)$$

Then the operation  $\mathcal{X}$  can be implemented efficiently: it simply maps  $|\vec{k}\rangle|0,\vec{0}\rangle \mapsto |\vec{k}\rangle|a,\vec{\theta}\rangle$ , where  $a$  and  $\vec{\theta}$  denote the set  $S_{a,\vec{\theta}}$  that contains  $\vec{k}$ .

Unfortunately, these window functions are sharply discontinuous, so the resulting curvelets are not very well-localized in space. This makes them poorly suited for the applications proposed in this paper (recall that the results of Sections 3, 4 and 5 required window functions that were  $C^1$ -smooth).

Smooth window functions are more challenging to implement, because the supports of the functions  $\chi_{a,\vec{\theta}}(\vec{k})$  necessarily overlap. Thus, at a given point  $\vec{k}$ , the operation  $\mathcal{X}$  must create a superposition of many values of  $a$  and  $\vec{\theta}$ . These superpositions can be complicated: for instance, if we imagine that the tiling of frequency space looks (locally) like an array of  $n$ -dimensional cubes, then a significant amount of volume lies near the corners of the cubes, and each corner point touches  $2^n$

different cubes, so we would have to prepare superpositions of  $2^n$  different values of  $a$  and  $\vec{\theta}$ . This seems impossible for many choices of the window functions.

However, the above example also suggests a solution to the problem. We can use spherical coordinates, which look locally like Cartesian coordinates (except at the poles). If we define the window functions to be products of simpler functions, each depending on a single variable, then we can prepare these superpositions efficiently. We now demonstrate this construction.

First, recall the definition of spherical coordinates in  $\mathbb{R}^n$ : we have  $(r, \phi_1, \dots, \phi_{n-1})$ , where  $r \in [0, \infty)$ ,  $\phi_1, \dots, \phi_{n-2} \in [0, \pi] \cup \{\text{"undef"}\}$ , and  $\phi_{n-1} \in (-\pi, \pi] \cup \{\text{"undef"}\}$ . We use the value “undef” to represent points on the “poles” of the sphere, e.g., if  $\phi_j = 0$  or  $\pi$ , then  $\phi_{j+1} = \dots = \phi_{n-1} = \text{"undef"}$  (a similar situation arises when  $r = 0$ ).

Cartesian coordinates are written in terms of spherical coordinates as follows:

$$x_1 = r \cos \phi_1 \quad (\text{or } 0 \text{ if undefined}), \quad (197)$$

$$x_j = r \sin \phi_1 \cdots \sin \phi_{j-1} \cos \phi_j \quad (\text{or } 0 \text{ if undefined}) \quad (j = 2, \dots, n-1), \quad (198)$$

$$x_n = r \sin \phi_1 \cdots \sin \phi_{n-1} \quad (\text{or } 0 \text{ if undefined}). \quad (199)$$

The reverse mapping is given by:

$$r = \sqrt{x_1^2 + \dots + x_n^2}, \quad (200)$$

$$\phi_1 = \arccos(x_1/r) \quad (\text{or “undef” if } r = 0), \quad (201)$$

$$\phi_j = \arccos(x_j/(r \sin \phi_1 \cdots \sin \phi_{j-1})) \quad (\text{or “undef” if } \phi_{j-1} \in \{0, \pi, \text{"undef"}\}) \quad (j = 2, \dots, n-2), \quad (202)$$

$$\phi_{n-1} = \text{sign}(x_n) \arccos(x_{n-1}/(r \sin \phi_1 \cdots \sin \phi_{n-2})) \quad (\text{or “undef” if } \phi_{n-2} \in \{0, \pi, \text{"undef"}\}). \quad (203)$$

Next we will define discrete values for the scale variable  $a$  and the direction variable  $\vec{\theta}$ . In the notation, it will be convenient to represent the scale variable by  $s$  instead of  $a$ , where

$$a = 2^{-s}. \quad (204)$$

We will then define window functions  $\chi_{s,\vec{\theta}}(\vec{k})$ . These will be products of radial and angular components (we write  $\vec{k} = (r, \vec{\phi})$  using spherical coordinates):

$$\chi_{s,\vec{\theta}}(\vec{k}) = w_s(\lambda r) v_{s,\vec{\theta}}(\vec{\phi}). \quad (205)$$

Here,  $\lambda$  is a parameter that sets the radial scaling. For future use, we define the function  $c : [0, \infty) \rightarrow \mathbb{R}$ ,

$$c(x) = \begin{cases} \cos x, & 0 \leq x \leq \pi/2, \\ 0, & x > \pi/2. \end{cases} \quad (206)$$

We begin with the scale variable  $a = 2^{-s}$ . We fix the cutoff values  $s_{min}, s_{max} \in \mathbb{Z}$ , where  $1 \leq s_{min} \leq s_{max}$ . Then we let  $s \in \{s_{min}, s_{min} + 1, \dots, s_{max}\} \cup \{\text{"coarse"}, \text{"fine"}\}$ .

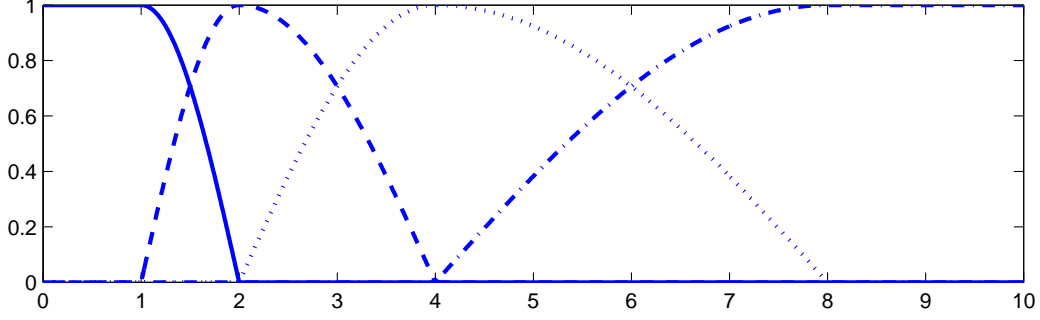


Figure 2: A family of radial window functions:  $w^{\text{coarse}}$ ,  $w_1$ ,  $w_2$  and  $w^{\text{fine}}$ .

We define radial window functions  $w_s(r)$  as follows:

$$w_s(r) = \begin{cases} c(\frac{\pi}{2}(2^s - r)/2^{s-1}), & 2^{s-1} \leq r \leq 2^s, \\ c(\frac{\pi}{2}(r - 2^s)/2^s), & 2^s \leq r \leq 2^{s+1}, \\ 0, & \text{otherwise,} \end{cases} \quad (207)$$

$$w^{\text{coarse}}(r) = \begin{cases} 1, & 0 \leq r \leq 2^{s_{\min}-1}, \\ c(\frac{\pi}{2}(r - 2^{s_{\min}-1})/2^{s_{\min}-1}), & 2^{s_{\min}-1} \leq r \leq 2^{s_{\min}}, \\ 0, & r \geq 2^{s_{\min}}, \end{cases} \quad (208)$$

$$w^{\text{fine}}(r) = \begin{cases} 0, & 0 \leq r \leq 2^{s_{\max}}, \\ c(\frac{\pi}{2}(2^{s_{\max}+1} - r)/2^{s_{\max}}), & 2^{s_{\max}} \leq r \leq 2^{s_{\max}+1}, \\ 1, & r \geq 2^{s_{\max}+1}. \end{cases} \quad (209)$$

An example is shown in Figure 2. It is easy to check that

$$\sum_s w_s(r)^2 = 1 \quad (\forall r \geq 0). \quad (210)$$

(Note that at any given point  $r$ , at most two of the functions  $w_s(r)$  are nonzero.)

We let the direction variable  $\vec{\theta}$  take on values in the set  $G_s(S^{n-1})$ . (Assume for the time being that  $s \notin \{\text{coarse}, \text{fine}\}$ ; we will handle those special cases later.) The set  $G_s(S^{n-1})$  contains grid points on the sphere  $S^{n-1}$ , defined using spherical coordinates, with angular spacing  $\pi/2^{\lceil s/2 \rceil} \approx \pi/\sqrt{2^s} = \pi\sqrt{a}$ . This set is defined recursively:

$$G_s(S^1) = \{\pi t/2^{\lceil s/2 \rceil} \mid t \in \mathbb{Z}, 0 \leq t \leq 2 \cdot 2^{\lceil s/2 \rceil} - 1\}, \quad (211)$$

$$G_s(S^k) = \{\pi t/2^{\lceil s/2 \rceil} \mid t \in \mathbb{Z}, 1 \leq t \leq 2^{\lceil s/2 \rceil} - 1\} \times G_s(S^{k-1}) \cup \{0, \pi\} \times \{\text{undef}\} \quad (k \geq 2). \quad (212)$$

(See Figure 3 for an example.)

We define angular window functions  $v_{s,\vec{\theta}}(\vec{\phi})$  as follows:

$$v_{s,\vec{\theta}}(\vec{\phi}) = \prod_{j=1}^{n-1} u_{s,\theta_j}(\phi_j), \quad (213)$$

where

$$u_{s,\theta_j}(\phi_j) = c(2^{\lceil s/2 \rceil} |\phi_j - \theta_j|/2). \quad (214)$$

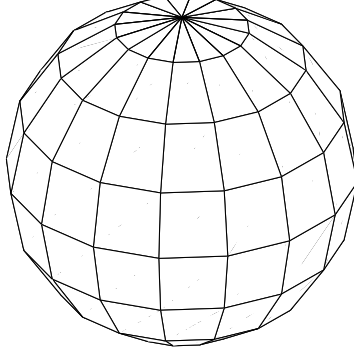


Figure 3:  $G_6(S^2)$ , the set of grid points on the sphere  $S^2 \subset \mathbb{R}^3$ , with angular spacing  $\pi/8$ .

This requires some further explanation.

Intuitively,  $u_{s,\theta_j}(\phi_j)$  is a one-dimensional “bump” function centered around  $\theta_j$ , and  $v_{s,\vec{\theta}}(\vec{\phi})$  is a product of these functions.

Note that, for the first  $n - 2$  coordinates  $j = 1, \dots, n - 2$ ,  $u_{s,\theta_j}(\phi_j)$  is defined on the interval  $[0, \pi]$ , whereas for the last coordinate  $j = n - 1$ ,  $u_{s,\theta_j}(\phi_j)$  is defined on the circle  $(-\pi, \pi]$ ; in this latter case, we interpret  $|\phi_j - \theta_j|$  as the shortest-path distance around the circle.

Also, in the definition of  $v_{s,\vec{\theta}}(\vec{\phi})$ , we simply omit those factors that have  $\theta_j = \text{“undef”}$  or  $\phi_j = \text{“undef”}$ . We claim that this is a natural thing to do, in that it yields a simple geometric picture. Intuitively,  $\theta_j = \text{“undef”}$  means that  $\vec{\theta}$  is located on a pole of the sphere, with some coordinate  $\theta_i$  ( $i < j$ ) equal to 0 or  $\pi$ . Then this construction produces a bump function that covers a circular region around the pole, and so does not depend on  $\phi_j$ . On the other hand, if  $\phi_j = \text{“undef”}$ , then  $\vec{\phi}$  is located on a pole of the sphere, with some coordinate  $\phi_i$  ( $i < j$ ) equal to 0 or  $\pi$ . If  $\theta_i \neq \phi_i$ , then  $u_{s,\theta_i}(\phi_i) = 0$ , hence  $v_{s,\vec{\theta}}(\vec{\phi}) = 0$ , independent of  $\phi_j$ . If  $\theta_i = \phi_i$ , then  $\vec{\theta}$  is located on a pole, hence  $v_{s,\vec{\theta}}(\vec{\phi})$  does not depend on  $\phi_j$ .

Note that at least the first ( $j = 1$ ) factor will always be defined, since  $\theta_1$  is always defined whenever  $\vec{\theta} \in G_s(S^{n-1})$ , and  $\phi_1$  is always defined whenever  $\vec{k} \neq \vec{0}$  (we can ignore the case of  $\vec{k} = \vec{0}$ , because it is relevant only when  $s = \text{“coarse”}$ , in which case we will not use these angular windows).

Finally, in the special case where  $s = \text{“coarse”}$  or  $\text{“fine”}$ , we do not resolve any directions  $\vec{\theta}$ . Instead, we fix  $\vec{\theta} = \text{“undef”}$ , and we define the angular window to be trivial,  $v_{s,\vec{\theta}}(\vec{\phi}) = 1$ .

We now show how to perform the operation  $\mathcal{X}$  that maps

$$|\vec{k}\rangle|0, \vec{0}\rangle \mapsto |\vec{k}\rangle \sum_{s, \vec{\theta}} \chi_{s, \vec{\theta}}(\vec{k}) |s, \vec{\theta}\rangle. \quad (215)$$

We will do this by converting  $\vec{k}$  to spherical coordinates  $(r, \vec{\phi})$ , performing an operation  $\mathcal{X}'$  that

creates the superposition over  $s$  and  $\vec{\theta}$ , then converting back to Cartesian coordinates:

$$\begin{aligned}
|\vec{k}\rangle|0, \vec{0}\rangle|0, \vec{0}\rangle &\mapsto |\vec{k}\rangle|r, \vec{\phi}\rangle|0, \vec{0}\rangle \\
&\mapsto |\vec{k}\rangle|r, \vec{\phi}\rangle \sum_{s, \vec{\theta}} w_s(\lambda r) v_{s, \vec{\theta}}(\vec{\phi}) |s, \vec{\theta}\rangle \\
&= |\vec{k}\rangle|r, \vec{\phi}\rangle \sum_{s, \vec{\theta}} \chi_{s, \vec{\theta}}(\vec{k}) |s, \vec{\theta}\rangle \\
&\mapsto |\vec{k}\rangle|0, \vec{0}\rangle \sum_{s, \vec{\theta}} \chi_{s, \vec{\theta}}(\vec{k}) |s, \vec{\theta}\rangle.
\end{aligned} \tag{216}$$

The operation  $\mathcal{X}'$  is implemented recursively, acting on the variables  $r, \phi_1, \dots, \phi_{n-1}$  one at a time:

```
// The “s” register
If  $\lambda r < 2^{s_{min}}$ , then let  $s_1 = \text{“coarse”}$  and  $s_2 = s_{min}$ .
Else if  $\lambda r > 2^{s_{max}}$ , then let  $s_1 = s_{max}$  and  $s_2 = \text{“fine”}$ .
Else, let  $s_1 = \lfloor \lg(\lambda r) \rfloor$  and  $s_2 = \lceil \lg(\lambda r) \rceil$ .
If  $s_1 = s_2$ , then set the “s” register to  $|s_1\rangle$ .
Else, set the “s” register to  $w_{s_1}(\lambda r)|s_1\rangle + w_{s_2}(\lambda r)|s_2\rangle$ .

// The “ $\theta_1$ ” register
If  $s \in \{\text{“coarse”}, \text{“fine”}\}$ , then set the “ $\theta_1$ ” register to  $|\text{“undef”}\rangle$ .
Else, begin:
  Let  $\tau_1 = (\pi/2^s)\lfloor \phi_1(2^s/\pi) \rfloor$  and  $\tau_2 = (\pi/2^s)\lceil \phi_1(2^s/\pi) \rceil$ .
  If  $\tau_1 = \tau_2$ , then set the “ $\theta_1$ ” register to  $|\tau_1\rangle$ .
  Else, set the “ $\theta_1$ ” register to  $u_{s, \tau_1}(\phi_1)|\tau_1\rangle + u_{s, \tau_2}(\phi_1)|\tau_2\rangle$ .
  // Note,  $\phi_1 \neq \text{“undef”}$ , since otherwise we would have  $s = \text{“coarse”}$ 
  // Note, if  $\phi_1 \in \{0, \pi\}$ , then  $\tau_1 = \tau_2$ , hence  $\theta_1 \in \{0, \pi\}$ 
End.
Recurse on the “ $\theta_2$ ” register.
```

```
// The “ $\theta_\ell$ ” register, for  $\ell = 2, \dots, n-2$ 
If  $\theta_{\ell-1} \in \{0, \pi, \text{“undef”}\}$ , then set the “ $\theta_\ell$ ” register to  $|\text{“undef”}\rangle$ .
Else, begin:
  Let  $\tau_1 = (\pi/2^s)\lfloor \phi_\ell(2^s/\pi) \rfloor$  and  $\tau_2 = (\pi/2^s)\lceil \phi_\ell(2^s/\pi) \rceil$ .
  If  $\tau_1 = \tau_2$ , then set the “ $\theta_\ell$ ” register to  $|\tau_1\rangle$ .
  Else, set the “ $\theta_\ell$ ” register to  $u_{s, \tau_1}(\phi_\ell)|\tau_1\rangle + u_{s, \tau_2}(\phi_\ell)|\tau_2\rangle$ .
  // Note,  $\phi_\ell \neq \text{“undef”}$ , since otherwise we would have, in some previous iteration  $k$ ,
  //  $\phi_k \in \{0, \pi\}$ , hence  $\theta_k \in \{0, \pi\}$ , and  $\theta_{\ell-1} \in \{0, \pi, \text{“undef”}\}$ 
  // Note, if  $\phi_\ell \in \{0, \pi\}$ , then  $\tau_1 = \tau_2$ , hence  $\theta_\ell \in \{0, \pi\}$ 
End.
Recurse on the “ $\theta_{\ell+1}$ ” register.
```

```
// The “ $\theta_{n-1}$ ” register
If  $\theta_{n-2} \in \{0, \pi, \text{“undef”}\}$ , then set the “ $\theta_{n-1}$ ” register to  $|\text{“undef”}\rangle$ .
Else, begin:
  Let  $\tau_1 = (\pi/2^s)\lfloor \phi_{n-1}(2^s/\pi) \rfloor$  and  $\tau_2 = (\pi/2^s)\lceil \phi_{n-1}(2^s/\pi) \rceil$ .
  If  $\tau_1 = \tau_2$ , then set the “ $\theta_{n-1}$ ” register to  $|\tau_1\rangle$ .
```

Else, set the “ $\theta_{n-1}$ ” register to  $\tilde{u}_{s,\tau_1}(\phi_{n-1})|\tau_1\rangle + \tilde{u}_{s,\tau_2}(\phi_{n-1})|\tau_2\rangle$ .  
 // Note,  $\phi_{n-1} \neq \text{“undef”}$ , since otherwise we would have, in some previous iteration  $k$ ,  
 //  $\phi_k \in \{0, \pi\}$ , hence  $\theta_k \in \{0, \pi\}$ , and  $\theta_{n-2} \in \{0, \pi, \text{“undef”}\}$   
 End.

This construction yields a fast quantum curvelet transform using smooth window functions. Note that we can carry out this construction using other choices of the function  $c(x)$ , which lead to different window functions  $\chi_{s,\vec{\theta}}(\vec{k})$ . We only need  $c(x)$  to satisfy the identity  $c(x)^2 + c(\frac{\pi}{2} - x)^2 = 1$  (for all  $0 \leq x \leq \pi/2$ ).

For instance, we can define

$$c(x) = \begin{cases} \cos(h(x)), & 0 \leq x \leq \pi/2, \\ 0, & x > \pi/2, \end{cases} \quad (217)$$

where  $h(x)$  is any increasing function that satisfies  $h(0) = 0$ ,  $h(\frac{\pi}{2}) = \frac{\pi}{2}$ , and  $h(\frac{\pi}{2} - x) = \frac{\pi}{2} - h(x)$  (for all  $0 \leq x \leq \pi/2$ ).

In particular, if we set  $h(x) = \frac{\pi}{2} \sin^2 x$ , then the resulting function  $c(x)$  is  $C^1$ -smooth. Thus we get window functions  $\chi_{s,\vec{\theta}}(\vec{k})$  that are  $C^1$ -smooth, and are qualitatively similar to the ones used in Sections 3-5 of this paper.

## F Quantum Algorithms using the Curvelet Transform

### F.1 Single-shot measurement of a quantum-sample state

Here we analyze a “continuous” analogue of our single-shot measurement procedure for finding the center of a ball.

First, we claim that  $a \leq \eta$  with probability  $\geq \Omega(\nu^2)$ . By Theorem 3, we have

$$\Pr[a \leq \eta] \geq (0.19)\eta(1 - \frac{1}{n}). \quad (218)$$

Note that  $\eta \geq \Omega(\nu^2)$ , which follows from the definition of  $\eta$  and the fact that  $\mu = \nu\beta \geq \nu\tilde{\beta}/2$ . This shows the claim.

From this point on, all probabilities are conditioned on having  $a \leq \eta$ .

Without loss of generality, assume  $\vec{c} = \vec{0}$ . The algorithm succeeds when it outputs a point close to  $\vec{0}$ . Let  $\Pi_1$  be the projector onto the subspace orthogonal to  $\vec{\theta}$ , and let  $\Pi_2$  be the projector onto the direction  $\vec{\theta}$ . We will show that, with constant probability,  $|\Pi_1\vec{b}|$  is small and  $|\Pi_2\vec{b}|$  is not too large.

Let  $X = |\Pi_1\vec{b}|^2$ ,  $\mu_X = \mathbb{E}(X)$ , and let  $Y = |\Pi_2\vec{b}|^2$ ,  $\mu_Y = \mathbb{E}(Y)$ . By Markov’s inequality,

$$\Pr[X \geq 3\mu_X] \leq \frac{1}{3}, \quad \Pr[Y \geq 3\mu_Y] \leq \frac{1}{3}.$$

Then, the union bound implies:

$$\Pr[X \leq 3\mu_X \text{ and } Y \leq 3\mu_Y] \geq 1 - \Pr[X \geq 3\mu_X] - \Pr[Y \geq 3\mu_Y] \geq \frac{1}{3}.$$

So, with probability  $\geq 1/3$ , we have  $X \leq 3\mu_X$  and  $Y \leq 3\mu_Y$ .

We now rewrite this in terms of  $|\Pi_1\vec{b}|$  and  $|\Pi_2\vec{b}|$ . From Theorem 3, we know that

$$\mu_X \leq \eta\beta^2(14300 + (Q_1/n)), \quad \mu_Y \leq \beta^2(242 + (Q_2/n)),$$

for some constants  $Q_1$  and  $Q_2$ . So, we have

$$|\Pi_1 \vec{b}| \leq \sqrt{3}\sqrt{\eta}\beta\sqrt{14300 + (Q_1/n)}, \quad |\Pi_2 \vec{b}| \leq \sqrt{3}\beta\sqrt{242 + (Q_2/n)}. \quad (219)$$

This shows that  $|\Pi_1 \vec{b}|$  (the error orthogonal to  $\vec{\theta}$ ) is small. Indeed, substituting in our choice of  $\eta$ , and using the fact that  $\beta/\tilde{\beta} \leq 1$ , we see that

$$|\Pi_1 \vec{b}| \leq \frac{1}{\sqrt{2}}\mu. \quad (220)$$

However,  $|\Pi_2 \vec{b}|$  (the error parallel to  $\vec{\theta}$ ) is not so small. So the algorithm tries to guess this error and output a corrected point. It succeeds when

$$|\Pi_2 \vec{b} - u(\sqrt{3}\tilde{\beta}\sqrt{242 + (Q_2/n)})\vec{\theta}| \leq \frac{1}{\sqrt{2}}\mu. \quad (221)$$

Call this event  $E$ . The probability of  $E$  is the probability that a random point chosen uniformly from the interval  $\sqrt{3}\tilde{\beta}\sqrt{242 + (Q_2/n)} \cdot [-1, 1]$  lies within distance  $\frac{1}{\sqrt{2}}\mu$  of some fixed point in the (possibly smaller) interval  $\sqrt{3}\beta\sqrt{242 + (Q_2/n)} \cdot [-1, 1]$ . This probability is lower-bounded by

$$\Pr[E] \geq \frac{\frac{1}{\sqrt{2}}\mu}{2\sqrt{3}\tilde{\beta}\sqrt{242 + (Q_2/n)}} = \frac{1}{2\sqrt{6}\sqrt{242 + (Q_2/n)}} \frac{\mu}{\tilde{\beta}}, \quad (222)$$

which is  $\geq \Omega(\nu)$ , since  $\mu = \nu\beta \geq \nu\tilde{\beta}/2$ .

So the algorithm succeeds with probability  $\geq \Omega(\nu^3)$ . This proves Theorem 5.  $\square$

Finally, we argue that, when the grid  $G = (\sigma\mathbb{Z})^n \cap [-L, L]^n$  is chosen properly, the discrete algorithm will behave like the continuous one. Recall from Section 6 that we can approximate a function  $f$  on  $\mathbb{R}^n$  with a function  $f_2$  on  $G$ , provided that most of the probability mass of  $f$  lies within distance  $L$  of the origin, and most of the probability mass of  $\hat{f}$  lies within distance  $\frac{1}{2\sigma}$  of the origin. This holds for our algorithm, provided that:

$$R + \beta \leq L, \quad \frac{100}{\lambda\eta} \leq \frac{1}{2\sigma}. \quad (223)$$

(The first condition follows immediately, since  $f$  is supported on a ball. The second condition follows from the decay of  $\hat{f}$ ; details omitted.) These conditions hold whenever

$$L \geq R + \tilde{\beta}, \quad \sigma \leq \frac{\pi e}{600} \frac{\mu^2}{\tilde{\beta}} \frac{1}{14300n + Q_1}. \quad (224)$$

Also, we argue that the discretization of the “direction” variable  $\vec{\theta}$  will not introduce too much error in the output of the algorithm. Our algorithm constructs a line  $\ell = \{\vec{b} + \lambda\vec{\theta} \mid \lambda \in \mathbb{R}\}$ , and if  $\vec{\theta}$  were a continuous variable, this line would pass within distance  $O(\sqrt{a}\beta)$  of the center  $\vec{c}$ . Recall from Section 6 that the discrete curvelet transform resolves  $\vec{\theta}$  within error  $\pm\sqrt{a}$  in angular distance. Since  $\vec{b}$  lies at distance  $O(\beta)$  from the center  $\vec{c}$ , the error in  $\vec{\theta}$  can increase the distance from  $\ell$  to  $\vec{c}$  by at most  $O(\sqrt{a}\beta)$ . Thus the error in the output of the algorithm increases by at most a constant factor.

We can bound the running time of our algorithm as follows. Let  $M$  be the number of grid points along one direction. Then  $M = 2L/\sigma$ , and the running time is  $\leq \text{poly}(n, \log M)$ . Say we choose  $L$  and  $\sigma$  so that the above inequalities are tight (up to constant factors). Then  $M \leq O(R\tilde{\beta}n/\mu^2)$ , and the running time is  $\leq \text{poly}(n, \log R, \log \tilde{\beta}, \log \frac{1}{\mu})$ .



## F.2 Quantum algorithm for finding the center of a radial function

Here we analyze the “continuous” analogue of our algorithm for finding the center of a radial function.

First, consider what happens for each  $i \in \{1, 2\}$ .

We quantum-sample over a ball of radius  $R'$  around  $\vec{0}$ , then measure the value of  $f$ , and get a superposition over a shell of radius  $\beta^{(i)}$  around  $\vec{c}$ . If the shell is too large, it will not lie completely within our original ball, so we only get a fragment of the shell. However, if  $\beta^{(i)} \leq R' - R$ , then we are guaranteed to get a complete shell.

We claim that we observe  $\beta^{(i)} \leq R' - R$  with constant probability. To see this, write:

$$\begin{aligned} \Pr[\beta^{(i)} \leq R' - R] &= \frac{\text{volume of ball of radius } R' - R \text{ around } \vec{c}}{\text{volume of ball of radius } R' \text{ around } \vec{0}} \\ &= \frac{(R' - R)^n}{(R')^n} = (1 - \frac{1}{n})^n \geq e^{-1.2} > 0.30, \end{aligned} \quad (225)$$

using the fact that  $1 - x \geq e^{-(1.2)x}$  for all  $0 \leq x \leq 1/4$  (recall that we assumed  $n \geq 4$ ).

Our algorithm does not know the shell’s true radius  $\beta^{(i)}$ , so it uses  $\tilde{\beta}^{(i)} = R' - R$  as an estimate. We claim that  $\beta^{(i)} \leq \tilde{\beta}^{(i)} \leq (3/2)\beta^{(i)}$ , with constant probability. Observe that

$$\begin{aligned} \Pr[\beta^{(i)} < (2/3)(R' - R)] &= \frac{\text{volume of ball of radius } (2/3)(R' - R) \text{ around } \vec{c}}{\text{volume of ball of radius } R' \text{ around } \vec{0}} \\ &= \frac{((2/3)(R' - R))^n}{(R')^n} = (2/3)^n (1 - \frac{1}{n})^n. \end{aligned} \quad (226)$$

So

$$\begin{aligned} \Pr[(2/3)(R' - R) \leq \beta^{(i)} \leq R' - R] &= (1 - (\frac{2}{3})^n)(1 - \frac{1}{n})^n \\ &\geq (1 - 0.20)(0.30) = 0.24, \end{aligned} \quad (227)$$

using the fact that  $n \geq 4$ .

Next, we claim that we observe  $a^{(i)} \leq \eta^{(i)}$ , with constant probability. This follows from Theorem 4:

$$\Pr[a^{(i)} \leq \eta^{(i)}] > 0.045. \quad (228)$$

From this point on, we take probabilities conditioned on  $a^{(i)} \leq \eta^{(i)}$ .

Without loss of generality, assume  $\vec{c} = \vec{0}$ . Let  $\Pi_1^{(i)}$  be the projector onto the subspace orthogonal to  $\vec{\theta}^{(i)}$ , and let  $\Pi_2^{(i)}$  be the projector onto the direction  $\vec{\theta}^{(i)}$ .

We claim that  $|\Pi_1^{(i)} \vec{b}^{(i)}|$  is small, and  $|\Pi_2^{(i)} \vec{b}^{(i)}|$  is of order  $\beta$ , with constant probability. We use the same argument as in the previous section, together with Theorem 4. We define  $\varepsilon^{(i)} = \delta/\beta^{(i)}$ . We get that, with probability  $\geq 1/3$ ,

$$|\Pi_1^{(i)} \vec{b}^{(i)}| \leq \sqrt{3} \sqrt{(n-1)\varepsilon^{(i)}\beta^{(i)}} \sqrt{761 + (Q_1/n)}, \quad |\Pi_2^{(i)} \vec{b}^{(i)}| \leq \sqrt{3}\beta^{(i)} \sqrt{23 + (Q_2/n^2)}, \quad (229)$$

for some constants  $Q_1$  and  $Q_2$ .

Using the definition of  $\varepsilon^{(i)}$ , and the fact that  $\beta^{(i)} \leq R' - R = (n-1)R$ , this implies that

$$|\Pi_1^{(i)} \vec{b}^{(i)}| \leq \sqrt{3}(n-1)\sqrt{\delta R} \sqrt{761 + (Q_1/n)}, \quad |\Pi_2^{(i)} \vec{b}^{(i)}| \leq \sqrt{3}(n-1)R \sqrt{23 + (Q_2/n^2)}. \quad (230)$$

The algorithm carries out this procedure twice, for  $i = 1$  and  $2$ . With constant probability, this produces two lines,  $\ell_1 = \{\vec{b}^{(1)} + \lambda \vec{\theta}^{(1)} \mid \lambda \in \mathbb{R}\}$  and  $\ell_2 = \{\vec{b}^{(2)} + \lambda \vec{\theta}^{(2)} \mid \lambda \in \mathbb{R}\}$ , which both pass near

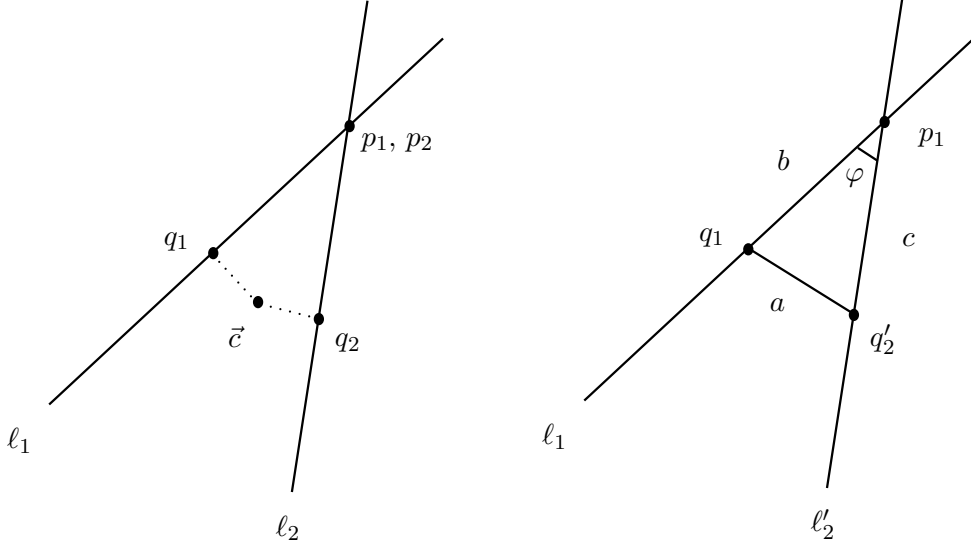


Figure 4:

the point  $\vec{c}$ . The algorithm then checks that these lines are nearly orthogonal, and if they are, it returns the point on  $\ell_1$  closest to  $\ell_2$ . A straightforward calculation shows that this point is given by  $\frac{-s+rt}{1-r^2}\vec{\theta}^{(1)} + \vec{b}^{(1)}$ .

First, we claim that the lines  $\ell_1$  and  $\ell_2$  are nearly orthogonal ( $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| \leq 3/4$ ) with at least constant probability.

We want to upper-bound the probability that  $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| > 3/4$ . Recall that these are independent random vectors, chosen uniformly from the unit sphere  $S^{n-1}$  in  $\mathbb{R}^n$ . It follows that

$$\Pr[|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| > 3/4] = \Pr[|x_1| > 3/4], \quad (231)$$

where  $\vec{x} = (x_1, \dots, x_n)$  is a random vector chosen uniformly from  $S^{n-1}$ . Note that  $E(\vec{x}) = \vec{0}$ , hence  $E(x_1) = 0$ ; also,  $E(|\vec{x}|^2) = 1$ , hence  $E(x_1^2) = 1/n$ . Then by Markov's inequality,

$$\Pr[|x_1| > 3/4] = \Pr[x_1^2 > 9/16] \leq \frac{16}{9} \frac{1}{n} \leq \frac{4}{9} \text{ (for } n \geq 4). \quad (232)$$

Thus, we observe  $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| \leq 3/4$ , with probability  $\geq 5/9$ . (This is a rather weak bound, especially when  $n$  is large, but it is adequate for our purposes. Actually, it is the case that  $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| \leq O(1/\sqrt{n})$ , with probability  $\geq \Omega(1)$ .)

Next, we claim that when  $\ell_1$  and  $\ell_2$  are nearly orthogonal ( $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| \leq 3/4$ ), the point on  $\ell_1$  closest to  $\ell_2$  (call it  $p_1$ ) is close to  $\vec{c}$ .

Let  $q_1$  be the point on  $\ell_1$  closest to  $\vec{c}$ , and let  $p_1$  be the point on  $\ell_1$  closest to  $\ell_2$ . Similarly, let  $q_2$  be the point on  $\ell_2$  closest to  $\vec{c}$ , and let  $p_2$  be the point on  $\ell_2$  closest to  $\ell_1$ . (See Figure 4.)

We know that  $q_1$  and  $q_2$  are both close to  $\vec{c}$ :  $|q_1 - \vec{c}| \leq \Delta$ , and  $|q_2 - \vec{c}| \leq \Delta$ , where  $\Delta = \sqrt{3(n-1)\sqrt{\delta R}\sqrt{761 + (Q_1/n)}}$ . Furthermore,  $p_1$  and  $p_2$  are close together:  $|p_1 - p_2| \leq 2\Delta$ .

Now suppose that  $p_1$  is far from  $\vec{c}$ :

$$|p_1 - \vec{c}| \geq 8\Delta. \quad (233)$$

From this we will derive a contradiction.

First, note that  $p_1$  is far from  $q_1$ :

$$|p_1 - q_1| \geq 7\Delta. \quad (234)$$

Consider the line  $\ell'_2 = \ell_2 + (p_1 - p_2)$ . Define another point  $q'_2 = q_2 + (p_1 - p_2)$ . The line  $\ell'_2$  is parallel to  $\ell_2$ , it intersects  $\ell_1$  at  $p_1$ , and it passes through  $q'_2$ . Note that  $q'_2$  is close to  $\vec{c}$ :  $|q'_2 - \vec{c}| \leq 3\Delta$ .

Note that  $p_1$  is far from  $q'_2$ :

$$|p_1 - q'_2| \geq 5\Delta. \quad (235)$$

Note that  $q_1$  and  $q'_2$  are close to each other:

$$|q_1 - q'_2| \leq 4\Delta. \quad (236)$$

We will use equations (234), (235) and (236) to show that the angle between  $\ell_1$  and  $\ell'_2$  is small. (See Figure 4.) We have  $a \leq 4\Delta$ ,  $b \geq 7\Delta$  and  $c \geq 5\Delta$ .

Using the law of cosines, and the fact that  $x + \frac{1}{x} \geq 2$  for all  $x > 0$ , we get:

$$\begin{aligned} |\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| &= \cos \varphi = \frac{c^2 + b^2 - a^2}{2bc} \\ &= \frac{1}{2} \left( \frac{c}{b} + \frac{b}{c} - \frac{a^2}{bc} \right) \\ &\geq 1 - \frac{a^2}{2bc} \\ &\geq 1 - \frac{(4\Delta)^2}{2(7\Delta)(5\Delta)} = 1 - \frac{8}{35} > \frac{3}{4}. \end{aligned} \quad (237)$$

This contradicts our assumption that  $\ell_1$  and  $\ell_2$  are nearly orthogonal.

So we conclude that  $p_1$  is close to  $\vec{c}$ , as desired:

$$|p_1 - \vec{c}| \leq 8\Delta = 8\sqrt{3}(n-1)\sqrt{\delta R}\sqrt{761 + (Q_1/n)}. \quad (238)$$

Finally, using our assumed upper bound on  $\delta$ , we get that:

$$|p_1 - \vec{c}| \leq \mu. \quad (239)$$

So the algorithm succeeds with constant probability. This proves Theorem 6.  $\square$

Finally, we argue that, when the grid  $G = (\sigma\mathbb{Z})^n \cap [-L, L]^n$  is chosen properly, the discrete algorithm will behave like the continuous one. Recall from Section 6 that we can approximate a function  $f$  on  $\mathbb{R}^n$  with a function  $f_2$  on  $G$ , provided that most of the probability mass of  $f$  lies within distance  $L$  of the origin, and most of the probability mass of  $\hat{f}$  lies within distance  $\frac{1}{2\sigma}$  of the origin. This holds for our algorithm, provided that:

$$R' \leq L, \quad \frac{100}{\delta} \leq \frac{1}{2\sigma}. \quad (240)$$

(The first condition follows immediately, since  $f$  is supported on a ball. The second condition follows from the decay of  $\hat{f}$ ; details omitted.)

Also, we argue that the discretization of the “direction” variable  $\vec{\theta}$  will not introduce too much error in the output of the algorithm. The key part of our algorithm involves constructing a line  $\ell = \{\vec{b} + \lambda\vec{\theta} \mid \lambda \in \mathbb{R}\}$ ; and if  $\vec{\theta}$  were a continuous variable, this line would pass within distance  $O(\sqrt{a}\beta)$  of the center  $\vec{c}$ . Recall from Section 6 that the discrete curvelet transform resolves  $\vec{\theta}$  within

error  $\pm\sqrt{a}$  in angular distance. Since  $\vec{b}$  lies at distance  $O(\beta)$  from the center  $\vec{c}$ , the error in  $\vec{\theta}$  can increase the distance from  $\ell$  to  $\vec{c}$  by at most  $O(\sqrt{a}\beta)$ . Thus the error in the output of the algorithm increases by at most a constant factor.

We can bound the running time of our algorithm as follows. Say we choose  $L$  and  $\sigma$  so that the above inequalities are tight (up to constant factors). Also, suppose  $\delta$  satisfies equation (??) exactly up to constant factors. Let  $M$  be the number of grid points along one direction. Then  $M = 2L/\sigma \leq O(R^2 n^3 / \mu^2)$ , and the running time is  $\leq \text{poly}(n, \log M) \leq \text{poly}(n, \log R, \log \frac{1}{\mu})$ .

### F.3 Classical lower bound

We will prove a lower bound on classical randomized algorithms for finding the center of a radial function. Recall that an instance of the problem is specified by an oracle  $f$  and parameters  $n$ ,  $R$ ,  $\delta$  and  $\mu$ . Let us define  $m = \lg(R/\mu)$ . Consider algorithms that query points within some finite subset  $D \subset \mathbb{R}^n$ . We will show that  $\Omega(nm/\lg(nm))$  queries are needed to solve this problem.

**Theorem 8** *The following holds for any parameters  $n$ ,  $R$ ,  $\delta$  and  $\mu$ , and for any finite subset  $D \subset \mathbb{R}^n$ . Let us define  $m = \lg(R/\mu)$ . For any classical randomized algorithm that queries points within the set  $D$  and uses at most  $(\frac{1}{2}nm)/\lg(\frac{1}{2}nm)$  oracle queries, there exists a problem instance (depending on  $D$ , and having the specified parameters  $n$ ,  $R$ ,  $\delta$  and  $\mu$ ) that causes the algorithm to fail with probability at least  $1 - 2^{-nm/2}$ .*

The assumption that the algorithm queries points within the set  $D$  can be understood intuitively as follows. We are assuming that the algorithm follows some (arbitrary) convention for how it describes points in  $\mathbb{R}^n$  when it queries the oracle. We let  $D \subset \mathbb{R}^n$  be the set of points that can be queried, and note that  $D$  must be finite: if the algorithm runs in time  $T$ , then clearly  $|D| \leq 2^T$ .

Note that this assumption does not weaken our lower bound. The assumption is needed because the construction of the hard instance depends on  $D$ ; however, the actual lower bound (i.e., the number of oracle queries and the resulting probability of success) is independent of  $D$ .

Proof: We will use the following version of Yao's minimax lemma, due to Rademacher and Vempala [24]:

**Lemma 9** *Let  $\mathcal{I}$  be a set of problem instances and  $\mathcal{A}$  be a set of deterministic algorithms. For any probability measure  $\pi$  over  $\mathcal{I}$ , and any probability measure  $\nu$  over  $\mathcal{A}$ , we have that*

$$\inf_{A \in \mathcal{A}} \Pr_{I \sim \pi(\mathcal{I})} [\text{algorithm } A \text{ fails on instance } I] \leq \sup_{I \in \mathcal{I}} \Pr_{A \sim \nu(\mathcal{A})} [\text{algorithm } A \text{ fails on instance } I]. \quad (241)$$

We will use the following approach. First we will fix a probability distribution  $\pi$  over instances, and prove that the best deterministic algorithm still fails with high probability. Note that a randomized algorithm is simply a probability distribution over deterministic algorithms; so the minimax lemma implies that for any randomized algorithm  $\nu$ , there exists an instance that causes the algorithm to fail with high probability.

We let  $\mathcal{A}$  be the set of deterministic algorithms that query points within the subset  $D$  and make at most  $\ell$  queries, for some  $\ell$  to be specified later.

We let  $\mathcal{I}$  be the set of problem instances, such that  $n$ ,  $R$ ,  $\delta$  and  $\mu$  are fixed, but  $f$  and  $\vec{c}$  can vary. That is, we fix the dimension  $n$ , the radius  $R$  of the ball in which the center lies, the thickness  $\delta$  of the spherical shells, and the desired accuracy  $\mu$ . We define  $m = \lg(R/\mu)$ , which is fixed. However, the values returned by the radial function  $f$ , and the location of its center  $\vec{c}$ , are arbitrary.

We now fix a distribution  $\pi$  on problem instances. Random instances according to this distribution are constructed as follows. First, we choose the center point  $\vec{c}$  uniformly at random from the ball of radius  $R$  around the origin. Let  $f$  be of the form

$$f(\vec{x}) = [s_k \text{ if } \vec{x} \in A_k], \text{ where } A_k = \{\vec{x} \mid k\delta \leq |\vec{x} - \vec{c}| < (k+1)\delta\}, \text{ for } k = 0, 1, 2, \dots \quad (242)$$

Next, we choose the values of  $f$  on the points in the set  $D$  (these are the points that the algorithm can query). Equivalently, we will choose the values of  $s_k$ , for those  $k$  such that  $A_k \cap D \neq \emptyset$ . Let  $K$  be the set of those  $k$ , and note that  $|K| \leq |D|$ . Let  $S = \{1, 2, \dots, |D|\}$ . Choose a random injective map  $\sigma : K \hookrightarrow S$ , and then set  $s_k = \sigma(k)$ .

Consider any deterministic algorithm  $A \in \mathcal{A}$ , and let  $U$  be the set of possible outputs of the algorithm. We claim that  $|U| \leq 2^{\ell \lg \ell}$ . To see this, note that the algorithm  $A$  can be described as a decision tree, where each node represents a query to the oracle  $f$ , and the algorithm chooses which branch to follow depending on the oracle's answer. We claim that after seeing the answer to its  $k$ 'th query, the algorithm can have at most  $k$  distinct branches. This is because, while the oracle can return many different values, they are meaningless except in cases where the oracle returns the same value as it did for a previous query. (Note that, for any permutation  $\sigma$  on the range of  $f$ , we can replace the oracle  $f$  with  $\sigma \circ f$ , to get a new instance of the problem that has the same desired solution. These two instances occur with equal probability under the distribution  $\pi$ .) So, when the algorithm receives the answer to its  $k$ 'th query, all it can do is compare that value to the answers to its previous queries. The number of branches is at most the number of distinct values that have been seen previously (which is at most  $k-1$ ), plus 1 (if the new value does not match any of the previous ones); thus the number of branches is at most  $k$ . Finally, note that the size of  $U$  is at most the number of leaves at the final level of the decision tree, i.e., after the  $\ell$ 'th query. So  $|U| \leq 1 \cdot 2 \cdot 3 \cdots \ell \leq \ell! = 2^{\ell \lg \ell}$ .

Thus we can upper-bound the probability that algorithm  $A$  succeeds on a random instance  $I \sim \pi(\mathcal{I})$ :

$$\begin{aligned} & \Pr_{I \sim \pi(\mathcal{I})} [\text{algorithm } A \text{ succeeds on instance } I] \\ & \leq \Pr_{\vec{c}} [\exists \vec{z} \in U \text{ s.t. } |\vec{z} - \vec{c}| \leq \mu] \leq \sum_{\vec{z} \in U} \Pr_{\vec{c}} [|\vec{z} - \vec{c}| \leq \mu] \\ & \leq \sum_{\vec{z} \in U} \frac{\text{volume of ball of radius } \mu}{\text{volume of ball of radius } R} = |U| \left(\frac{\mu}{R}\right)^n \\ & \leq 2^{\ell \lg \ell} 2^{-mn}. \end{aligned} \quad (243)$$

Now suppose that  $\ell$ , the number of queries, is at most  $(\frac{1}{2}nm)/\lg(\frac{1}{2}nm)$ . Then

$$\ell \lg \ell \leq \frac{\frac{1}{2}nm}{\lg(\frac{1}{2}nm)} (\lg(\frac{1}{2}nm) - \lg \lg(\frac{1}{2}nm)) \leq \frac{1}{2}nm \text{ (assuming } nm \geq 4). \quad (244)$$

So we have

$$\Pr_{I \sim \pi(\mathcal{I})} [\text{algorithm } A \text{ succeeds on instance } I] \leq 2^{-nm/2}, \quad (245)$$

and this holds for all algorithms  $A \in \mathcal{A}$ . So

$$\inf_{A \in \mathcal{A}} \Pr_{I \sim \pi(\mathcal{I})} [\text{algorithm } A \text{ fails on instance } I] \geq 1 - 2^{-nm/2}. \quad (246)$$

Now plug into the minimax lemma and the result follows.  $\square$

#### F.4 Finding the center through multiple iterations

Here we analyze the “continuous” analogue of our algorithm for finding the center of a radial function.

First, we analyze the procedure `OneRound()`. Consider what happens for each  $i \in \{1, 2\}$ .

We quantum-sample over a ball of radius  $R'$  around  $\vec{0}$ , then measure the value of  $f$ , and get a superposition over a shell of radius  $\beta^{(i)}$  around  $\vec{c}$ . If the shell is too large, it will not lie completely within our original ball, so we only get a fragment of the shell. However, if  $\beta^{(i)} \leq R' - R$ , then we are guaranteed to get a complete shell.

We claim that we observe  $\beta^{(i)} \leq R' - R$  with probability  $\geq 1 - O(\frac{1}{S})$ . To see this, write:

$$\begin{aligned} \Pr[\beta^{(i)} \leq R' - R] &= \frac{\text{volume of ball of radius } R' - R \text{ around } \vec{c}}{\text{volume of ball of radius } R' \text{ around } \vec{0}} \\ &= \frac{(R' - R)^n}{(R')^n} = (1 - \frac{1}{nS})^n \geq e^{-(1.2)/S} \geq 1 - \frac{1.2}{S}, \end{aligned} \quad (247)$$

using the fact that  $1 - x \geq e^{-(1.2)x}$  for all  $0 \leq x \leq 1/4$  (recall that we assumed  $n \geq 4$ ).

Our algorithm does not know the shell's true radius  $\beta^{(i)}$ , so it uses  $\tilde{\beta}^{(i)} = R' - R$  as an estimate. We claim that  $\beta^{(i)} \leq \tilde{\beta}^{(i)} \leq S\beta^{(i)}$ , with probability  $\geq 1 - O(\frac{1}{S})$ . Observe that

$$\begin{aligned} \Pr[\beta^{(i)} < (1/S)(R' - R)] &= \frac{\text{volume of ball of radius } (1/S)(R' - R) \text{ around } \vec{c}}{\text{volume of ball of radius } R' \text{ around } \vec{0}} \\ &= \frac{((1/S)(R' - R))^n}{(R')^n} = (1/S)^n (1 - \frac{1}{nS})^n. \end{aligned} \quad (248)$$

So

$$\begin{aligned} \Pr[(1/S)(R' - R) \leq \beta^{(i)} \leq R' - R] &= (1 - (\frac{1}{S})^n)(1 - \frac{1}{nS})^n \\ &\geq (1 - \frac{1}{S})(1 - \frac{1.2}{S}) > 1 - \frac{2.2}{S}. \end{aligned} \quad (249)$$

Next, we claim that we observe  $a^{(i)} \leq \eta^{(i)}$ , with constant probability. This follows from Theorem 4:

$$\Pr[a^{(i)} \leq \eta^{(i)}] > 0.045. \quad (250)$$

From this point on, we take probabilities conditioned on  $a^{(i)} \leq \eta^{(i)}$ .

Without loss of generality, assume  $\vec{c} = \vec{0}$ . Let  $\Pi_1^{(i)}$  be the projector onto the subspace orthogonal to  $\vec{\theta}^{(i)}$ , and let  $\Pi_2^{(i)}$  be the projector onto the direction  $\vec{\theta}^{(i)}$ .

We claim that  $|\Pi_1^{(i)} \vec{b}^{(i)}|$  is small, and  $|\Pi_2^{(i)} \vec{b}^{(i)}|$  is of order  $\beta$ , with probability  $\geq 1 - O(\frac{1}{S})$ . We use a similar argument as in the previous section, together with Theorem 4. We define  $\varepsilon^{(i)} = \delta/\beta^{(i)}$ . We get that, with probability  $\geq 1 - \frac{2}{S}$ ,

$$|\Pi_1^{(i)} \vec{b}^{(i)}| \leq \sqrt{S} \sqrt{(n-1)\varepsilon^{(i)}\beta^{(i)}} \sqrt{(507 + \frac{Q_1}{n}) \cdot S}, \quad |\Pi_2^{(i)} \vec{b}^{(i)}| \leq \sqrt{S}\beta^{(i)} \sqrt{23 + \frac{Q_2}{n^2}}, \quad (251)$$

for some constants  $Q_1$  and  $Q_2$ .

Using the definition of  $\varepsilon^{(i)}$ , and the fact that  $\beta^{(i)} \leq R' - R = (nS - 1)R < nSR$ , this implies that

$$|\Pi_1^{(i)} \vec{b}^{(i)}| \leq S^{3/2} n \sqrt{\delta R} \sqrt{507 + \frac{Q_1}{n}}, \quad |\Pi_2^{(i)} \vec{b}^{(i)}| \leq S^{3/2} n R \sqrt{23 + \frac{Q_2}{n^2}}. \quad (252)$$

The algorithm carries out this procedure twice, for  $i = 1$  and  $2$ . With probability  $\geq 1 - O(\frac{1}{S})$ , this produces two lines,  $\ell_1 = \{\vec{b}^{(1)} + \lambda \vec{\theta}^{(1)} \mid \lambda \in \mathbb{R}\}$  and  $\ell_2 = \{\vec{b}^{(2)} + \lambda \vec{\theta}^{(2)} \mid \lambda \in \mathbb{R}\}$ , which both pass

near the point  $\vec{c}$ . The algorithm then checks that these lines are nearly orthogonal, and if they are, it returns the point on  $\ell_1$  closest to  $\ell_2$ . A straightforward calculation shows that this point is given by  $\frac{-s+rt}{1-r^2}\vec{\theta}^{(1)} + \vec{b}^{(1)}$ .

First, we claim that the lines  $\ell_1$  and  $\ell_2$  are nearly orthogonal ( $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| \leq 3/4$ ) with probability  $\geq 5/9$ . This follows from the same argument as in the previous section.

Next, we claim that when  $\ell_1$  and  $\ell_2$  are nearly orthogonal ( $|\vec{\theta}^{(1)} \cdot \vec{\theta}^{(2)}| \leq 3/4$ ), the point on  $\ell_1$  closest to  $\ell_2$  (call it  $p_1$ ) is close to  $\vec{c}$ . Using the same argument as in the previous section, we conclude that:

$$|p_1 - \vec{c}| \leq 8S^{3/2}n\sqrt{\delta R}\sqrt{507 + \frac{Q_1}{n}}. \quad (253)$$

Finally, using our assumed upper bound on  $\delta$ , we get that:

$$|p_1 - \vec{c}| \leq \sqrt{R}\sqrt{\mu/2}. \quad (254)$$

In summary, we have shown that the procedure `OneRound()` has the following two properties:

(1) `OneRound()` returns a point  $\vec{q}$  (rather than “no answer”) with probability

$$\geq (0.045)^2 \cdot \frac{5}{9} > 0.0011, \quad (255)$$

and (2) when `OneRound()` returns a point  $\vec{q}$ , that point  $\vec{q}$  lies within distance  $\sqrt{R}\sqrt{\mu/2}$  of the center point  $\vec{c}$ , with probability

$$\geq \left(1 - \frac{2.2}{S}\right)\left(1 - \frac{2}{S}\right)^2 > \left(1 - \frac{4.2}{S}\right)^2 > 1 - \frac{8.4}{S}. \quad (256)$$

We now analyze the complete algorithm, consisting of multiple iterations.

First, let us upper-bound the number of iterations, assuming that every iteration is successful. Let  $k$  denote the number of iterations. We claim that  $k \leq \lceil \lg \lg \frac{2R}{\mu} \rceil$ .

Let  $R_i$  denote the value of  $R_{cur}$  following the  $i$ 'th iteration, so we have  $R_0 = R$  (when the algorithm starts),  $R_{i+1} = \sqrt{R_i}\sqrt{\mu/2}$  (the recurrence relation), and  $R_k \leq \mu$  (when the algorithm finishes).

Let us define  $\tilde{R}_i = \frac{R_i}{\mu/2}$ . Then we have  $\tilde{R}_0 = \frac{R}{\mu/2}$  (when the algorithm starts),  $\tilde{R}_{i+1} = \sqrt{\tilde{R}_i}$  (the recurrence relation), and  $\tilde{R}_k \leq 2$  (when the algorithm finishes).

It is easy to see that  $\tilde{R}_k = (\tilde{R}_0)^{(1/2)^k}$ , and a straightforward calculation shows that it suffices to set  $k = \lceil \lg \lg \frac{2R}{\mu} \rceil$ .

Next, we show that the algorithm succeeds (i.e., every iteration is successful) with constant probability. First, consider a single iteration. The algorithm makes  $n_{tries}$  attempts to run `OneRound()`, and it succeeds if at least one of those attempts returns a point  $\vec{q}$  that lies near the center. The probability that `OneRound()` returns “no answer” every time is  $\leq (0.9989)^{n_{tries}} = (0.9989)^{910 \lg S} \leq 1/S$ . So the probability that `OneRound()` returns a point  $\vec{q}$  at least once is  $\geq 1 - \frac{1}{S}$ . When this happens, the point  $\vec{q}$  is near the center with probability  $\geq 1 - \frac{8.4}{S}$ . So the iteration succeeds with overall probability  $\geq (1 - \frac{1}{S})(1 - \frac{8.4}{S}) > 1 - \frac{9.4}{S} = 1 - \frac{1}{n_{iter}}$ .

The algorithm makes  $n_{iter}$  iterations, and it succeeds if all of the iterations succeed. This occurs with probability  $\geq (1 - \frac{1}{n_{iter}})^{n_{iter}} \geq 1/e^2$ . (We assumed that  $R \geq 8\mu$ , so  $n_{iter} \geq 2$ , and we used the fact that  $1 - x \geq e^{-2x}$  for all  $0 \leq x \leq 1/2$ .) So the algorithm succeeds with constant probability.

Finally, note that the number of oracle queries is  $2n_{tries}n_{iter} \leq O(\lg \lg \frac{2R}{\mu} \lg \lg \lg \frac{2R}{\mu})$ . This proves Theorem 7.  $\square$

Finally, we argue that, when the grid  $G = (\sigma\mathbb{Z})^n \cap [-L, L]^n$  is chosen properly, the discrete algorithm will behave like the continuous one. Recall from Section 6 that we can approximate a function  $f$  on  $\mathbb{R}^n$  with a function  $f_2$  on  $G$ , provided that most of the probability mass of  $f$  lies within distance  $L$  of the origin, and most of the probability mass of  $\hat{f}$  lies within distance  $\frac{1}{2\sigma}$  of the origin. This holds for our algorithm, provided that:

$$R' \leq L, \quad \frac{100}{\delta} \leq \frac{1}{2\sigma}. \quad (257)$$

(The first condition follows immediately, since  $f$  is supported on a ball. The second condition follows from the decay of  $\hat{f}$ ; details omitted.)

Also, we argue that the discretization of the “direction” variable  $\vec{\theta}$  will not introduce too much error in the output of the algorithm. The key part of our algorithm involves constructing a line  $\ell = \{\vec{b} + \lambda\vec{\theta} \mid \lambda \in \mathbb{R}\}$ ; and if  $\vec{\theta}$  were a continuous variable, this line would pass within distance  $O(\sqrt{a}\beta)$  of the center  $\vec{c}$ . Recall from Section 6 that the discrete curvelet transform resolves  $\vec{\theta}$  within error  $\pm\sqrt{a}$  in angular distance. Since  $\vec{b}$  lies at distance  $O(\beta)$  from the center  $\vec{c}$ , the error in  $\vec{\theta}$  can increase the distance from  $\ell$  to  $\vec{c}$  by at most  $O(\sqrt{a}\beta)$ . Thus the error in the output of the algorithm increases by at most a constant factor.

We can bound the running time of our algorithm as follows. Say we choose  $L$  and  $\sigma$  so that the above inequalities are tight (up to constant factors). Also, suppose  $\delta$  satisfies equation (??) exactly up to constant factors. Let  $M$  be the number of grid points along one direction. Then  $M = 2L/\sigma \leq O(R^2 n^3 / \mu^2)$ , and the running time is  $\leq \text{poly}(n, \log M) \leq \text{poly}(n, \log R, \log \frac{1}{\mu})$ .